Problems of Bayesian trait reconstruction on linguistic phylogenies in evolutionary anthropology

Étienne Fortier-Dubois

Master report
under the supervision of Joe Henrich, Harvard University,
and Franjo Weissing, University of Groningen

August 2016

Abstract

The parallel between biological evolution and cultural evolution allows, with some caveats, the use of phylogenetic comparative methods in anthropology, including ancestral trait reconstruction. The controversy over such methods requires additional evaluation. I used the program BayesTraits, together with phylogenetic trees built from linguistic and historical data, to perform Bayesian reconstructions of world religions and ethnological traits in cultures of the Indo-European language family. By excluding certain taxa from an analysis, performing a reconstruction on different tree sets, and comparing the reconstruction with the known historical record, I highlight some issues of the method that must be taken into account in future studies.

Contents

Introduction 2
  Problems of Bayesian trait reconstruction 3
  Previous studies and aims 4
Materials and Methods 5
Results 9
  Marriage type 9
  World religions 11
  Other traits 14
Discussion 18
  Incorporation of history 18
  Correctness of the tree 20
  Horizontal transmission 21
  Conclusion 22
Appendix A: Table of dates of Christianization 27
Appendix B: Reconstruction of some traits from the *Ethnographic Atlas* 29

Introduction

The parallel between biological evolution and cultural evolution has been thoroughly documented, analyzed, and criticized (Mace et al., 1994; Atkinson and Gray, 2005; Mace and Holden, 2005). Cultural units such as languages, traditional practices, and technologies share several common points with genetic information; and cultures, understood as "behavioural traditions that are transmitted by social learning" (Mace and Holden, 2005), with biological taxa. Although the analogy is far from perfect (Tëmkin and Eldredge, 2007; Gray et al., 2007; Greenhill et al., 2009), it has proven sufficient to allow the adoption of several methods from evolutionary biology by practitioners of the social sciences. A notable example is in linguistics, where the use of phylogenetic trees is now widespread to describe the relationships between languages. In fact, tree-thinking developed independently, and in ways very alike, in the study of both languages and living forms (Atkinson and Gray, 2005).

In biology, phylogenetic comparative methods have been developed since the 1980s as a solution to the problem of statistical non-independence of species (Cooper et al., 2016). This problem has also long been recognized in anthropology, where it is known as "Galton's problem" from the work by Francis Galton in the 1880s (Mace et al., 1994). Researchers dealing with cross-cultural comparison have retorted to various solutions to Galton's problem, such as creating a "standard cross-cultural sample" (Murdock and White, 1969), but phylogeny has stood out as perhaps the most robust way to perform comparisons while taking into account the non-independence of cultures. Several types of questions can be asked using phylogenetic comparative methods. Within the realm of cultural evolution, Gray et al. (2007) identify six: 1) searching for homelands; 2) identifying sequences of cultural change; 3) dating divergence events; 4) calculating rates of cultural change; 5) testing adaptation hypotheses; and 6) reconstructing ancestral states. In this paper, the focus is on the latter.

Several methods can be used to perform phylogenetic analyses, such as maximum parsimony, but in recent times, Bayesian statistical procedures have become of increased in the field (Pagel et al., 2004). Bayesian phylogenetic comparative methods are not, however, exempt from controversy (Gray et al., 2007; Greenhill et al., 2009). In a general sense, just like other statistical methods, they involve assumptions and biases that may invalidate the conclusions drawn from them (Cooper et al., 2016). Several limitations and problems also arise from their specific application to cultural evolution.

Problems of Bayesian trait reconstruction

The most apparent such problem may be horizontal transmission. Phylogenies carry the underlying assumption that traits are passed down 'vertically', that is, from an ancestor to its descendents. This assumption generally holds true in the realm of biological species, where vertical transmission of genetic material is the norm. Many (e.g. Gould, 1987; Gray et al., 2007; Tëmkin and Eldredge, 2007; Lewens, 2013; Testart, 2013) have pointed out that it is not the case in cultural evolution, however. Cultural practices may be transmitted both vertically or horizontally, and it is unclear which of these is the most important process (Collard et al., 2006), and what this means for uncovering the relationships between societies. As an example, a religion may be present in a culture either because its mother culture had it, or because it was introduced independently from ancestry.

As a result, relationships between cultures may be more reticulated than simple phylogenetic trees would suggest. Must phylogenetic inferences in cultural evolution, then, be rejected? Greenhill et al. (2009) suggest that they should not, based on an analysis in which they attempted to infer linguistic phylogenetic information in a model where there were "realistic levels of borrowing". They concluded that phylogenies are resistant to some degree of reticulation. Collard et al. (2006) reached a similar conclusion from fitting cultural and biological datasets on tree models and finding that they do not significantly differ. They point out that branching processes cannot be discounted in studies of cultural evolution, and that whether a phylogenetic model is appropriate or not depends on the dataset.

Another important obstacle in phylogenetic trait reconstruction is fundamental in nature: it requires a correct phylogenetic tree. Indeed, any reconstruction of an ancestral state is valid only insofar as the 'ancestor' truly exists, which requires the phylogeny to be an accurate representation of the evolution of the group. Bayesian phylogenetic inference, by incorporating inherent uncertainty in the resulting trees, partially resolve this issue (Pagel et al., 2004). But even so, the data may impose limitations. Thus, while languages can serve as a good proxy for cultural relatedness (Mace et al., 1994; Jordan et al., 2009), and while they generally agree with genetically-derived phylogenies (Cavalli-Sforza, 1997), they are not a perfect phylogenetic tool.

One problem is that relationships between the main families (Indo-European, Bantu, Sino-Tibetan, Finno-Ugric, etc.) are not only controversial, but may never be resolved due to the ancestors of "superfamilies" being so ancient that any features that could serve as comparison points have possibly disappeared due to normal linguistic evolution (Ringe, 1995; Pagel et al., 2013, although the latter argue that it is possible to find evidence for superfamilies in some ultraconserved linguistic features). This makes it difficult to compare societies from different language families and to perform worldwide analyses. Although Peoples et al. (2016) have attempted religious trait reconstruction on a worldwide tree, they did not provide details on how they obtained their tree. A second problem is that languages may not be the best indicator of the (biological or cultural) evolution of a group. As pointed out by Cavalli-Sforza (1997), a number of processes, like demic expansion, military conquest or economic control, can create patterns in the evolution of languages that are independent from other evolutionary patterns.

Lastly, phylogenetic comparative methods that aim at reconstructing history have attracted criticism for not taking into account known historical facts. For instance, Testart (2013) has criticized a study of dowry and bridewealth in the Indo-European space by Fortunato et al. (2006), pointing out that they failed to take into account extant historical knowledge despite their claim to have done so. By performing a more conventional historical analysis, Testart reached a conclusion opposed to theirs. This highlights the need for some sort of validation of phylogenetic methods when applied to human history.

Previous studies and aims

These issues have not prevented phylogenetic comparative methods from gaining a foothold in evolutionary anthropology. Several studies now provide examples of Bayesian trait reconstruction, either for its own sake or as a way of investigating the coevolution of cultural traits. In the Indo-European family, Fortunato et al. (2006) studied the ancestral state of marriage transfer (dowry, bridewealth) as mentioned above, and Fortunato later attempted a reconstruction of marriage type (monogamy or polygamy; Fortunato, 2011a) and post-marital residence (Fortunato, 2011b). Walker et al. (2011) have similarly studied marriage practices in worldwide hunter-gatherer societies, though using mitochondrial DNA rather than languages as a phylogenetic basis. Jordan et al. (2009) and Currie et al. (2010) have applied the method to the Austronesian language family, and reconstructed, respectively, the history of post-marital residence and that of political complexity in Southeast Asia and the Pacific. This in turn allowed studies of the coevolutionary relationship between political complexity and religion (Watts et al., 2015) as well as human sacrifice (Watts et al., 2016) in Austronesia. Other studies have focussed on various traits in the South American Tupi (Walker et al., 2012) and African Bantu (Opie et al., 2014) language families.

This wealth of phylogenetic studies of cultural trait reconstruction should not mask the fact that much of the criticism has not been properly addressed yet. In particular, there has been very little benchmarking done to evaluate the results of trait reconstruction against known historical data. I seeked to remedy this by reconstructing traits that have a well-known history in a well-documented linguistic space, the Indo-European family. I also examined the important effect that the inclusion or exclusion of some taxa may have on the reconstruction.

Materials and Methods

In order to perform phylogenetic trait reconstruction, I used the software BayesTraits V2 (Pagel and Meade, 2006; Meade and Pagel, 2014). More specifically, I used BayesTraits' MultiState program with the Markov chain Monte Carlo (MCMC) type of analysis. A detailed explanation of Bayesian phylogenetics will not occur here, but briefly, the MCMC algorithm is used to sample the space of all possible states in proportion to their likelihoods. A 'state', in this context, refers to a particular assignment of values to the parameters of interest (which are initially given 'prior' distributions). For instance, in a given state, a node in the tree will be reconstructed with a specific trait value; if a state containing this reconstruction is sampled 75% of the time, then we can interpret the reconstruction as having a 75% 'posterior' probability. For more details, the reader is referred especially to Pagel et al. (2004); see also Fortunato (2011a).

BayesTraits takes two elements of input: a tree (or a set of trees), and a file containing trait coding information. While creating trees from genetic, linguistic or material culture data is a major application of Bayesian phylogenetics (Huelsenbeck et al., 2001), I did not perform this here. Rather, I reused Bayesian tree samples from previous studies. In the Indo-European space, comprehensive tree samples have been created in Pagel et al. (2007)

and Bouckaert et al. (2012). The former consists of 751 trees and 87 languages, and the branch lengths are proportional to linguistic change; in the latter, which comprises 10,503 trees and 103 languages, the branch lengths are proportional to time in years. This incorporation of temporal information relies on the callibration of extinct languages and some internal nodes using historical information, and it has the advantage of allowing dates to be marked directly on the trees. This is why I used the Bouckaert et al. (2012) trees for most analyses, and used the Pagel et al. (2007) set only for comparison purposes. Both sets of trees were built from a vocabulary database by Dyen et al. (1992), although the more recent set incorporates a large number of other linguistic sources as well (Bouckaert et al., 2012, supplementary information). The temporal placement of extinct language tips reflects their earliest well-attested date from written texts (Ringe et al., 2002); the other tips are placed in the present time. The naming of the languages in the trees, as seen in figures and tables here, is from Dyen et al. (1992).

The other required element for trait reconstruction is trait coding. The sources depended on the trait, as described below.

Outside of BayesTraits, data preparation, processing and analysis took place within the R framework, version 3.3.0 (R Development Core Team, 2008). R was supplemented by two packages designed for phylogenetic work: ape 3.5 (Paradis et al., 2004), and phangorn 2.0.3 (Schliep, 2011). In R I performed such tasks as scaling the trees (BayesTraits works best when branch lengths are on the order of fractions of a unit, which is not the case if they represent years; Quentin D. Atkinson, personal communication), pruning them, and visualizing them.

Running a MCMC chain requires some parameters and priors to be set. To get a sense of what specifications were needed, I first replicated the study of marriage type by Fortunato (2011a). In this paper, the author seeked to determine whether the ancestors of Indo-European–speaking peoples were monogamous or polygamous based on the data on recent peoples from the *Ethnographic Atlas*. I used the same tree set (Pagel et al., 2007) and data, as well as the more recent tree set (Bouckaert et al., 2012) for comparison. BayesTraits attemps to model the evolution of a binary trait through the use of two parameters, which are the rates of change between trait states: from monogamy to polygamy (q_{MP}), and from polygamy to monogamy (q_{PM}). As Fortunato did, I used a Reverse-Jump MCMC (RJMCMC) approach (Pagel and Meade, 2006), which searches among models of trait evolution at the same time as it estimates the parameters. The models are distinguished into four classes based on whether the parameters are equal and whether they are positive or zero. The four model classes are: 1) $q_{MP} = q_{PM} > 0$; 2) $q_{MP} = 0 < q_{PM}$; 3) $q_{MP} > 0 = q_{PM}$; and 4) $q_{MP} \neq q_{PM}$ where $q_{MP}, q_{PM} > 0$. The RJMCMC option was assigned a hyperprior, which is a distribution from which are drawn values determining the priors, themselves distributions sampled to assign values to the parameters during the RJMCMC run. The use of hyperpriors is recommended by the BayesTraits manual (Meade and Pagel, 2014) because they reduce arbitrariness. In this case, the prior was an exponential distribution whose mean was sampled from a (hyperprior) uniform distribution between 0 and 10. As for other BayesTraits options, I ran the chain for 100 million iterations, sampling every 1000, and discarding the first million as a burn-in period in which convergence may not have been attained yet. Default options were used for the rest. Unless otherwise specified, these are the values used for all the reported MCMC analyses.

I then reconstructed the presence or absence of world religions (Christianity, Islam, and Hinduism) in the ancestors of Indo-European–speaking peoples. World religions lend themselves well for this analysis, because 1) the three mentioned are widespread among Indo-European peoples, but not ubiquitous; 2) is is easy to consider each as a binary trait, and it is generally straightforward to decide if a culture has a particular world religion as a dominant religion or not; and 3) their appearance in a culture can often be dated from historical documentation. Using the list of languages found in Bouckaert et al. (2012), I pruned the trees from languages that would refer to approximately the same culture as others (e.g. dialects of Sardinian were collpased into a single Sardinian taxon). I coded the remaining 87 language groups depending on whether the religion of interest was or is the main religion of that society, using general knowledge and encyclopedias (Lochtefeld, 2002; Waldman and Mason, 2006; Kurian, 2011; Martin, 2016). For the 20 extinct groups, the coding depended on the date at which they were converted (see below) as compared to their placement on the temporal axis of the tree. Uniquely, I coded the Roma people as ambiguous for Christianity and Islam since they include populations following both these faiths (Lewis, 1991).

Without any prior information other than the tree and the coding, I ran the trait reconstruction in BayesTraits. In order to compare the results of the Christianity reconstruction with known history, I then assigned a date representing conversion for each relevant group (I did not perform this for Islam and Hinduism). It is not straightforward to assign such a date, since the spread of a religion can be a slow, incremental process; thus the assigned dates must be regarded as approximative. They may represent, for instance, the first baptism of a ruler, or the period of activity of an important missionary. A key date is that of the Edict of Milan, in 313, which lead to acceptation of Christianity in the Roman Empire (Waldman and Mason, 2006, p. 658). It is the date used for most areas that were part of the empire at that time, even though the process of Christianization of such a large state as the Roman Empire was complex and can almost certainly not be summarized by a single year. The dates and sources for all groups are given in Appendix A.

For other traits, the source is the *Ethnographic Atlas*, published by Murdock in several installments in the journal *Ethnology*, summarized by Murdock (1967), and corrected by Gray (1999). The *Atlas* includes data on several variables for a large number of societies (usually from the first half of the twentieth century), but it does not cover all of them, especially in the European area. Thus, only a subset (at most 35) of the languages found in the tree samples could be associated with a culture in the *Atlas*. In addition, I mapped some of the languages to colonial European populations rather than to their homeland population, which were absent from the *Atlas*; this is the case for English (New England Americans), French (French Canadians) and part of the Portuguese coding (Brazilians). There is certainly some variation between Europeans and their daughter colonial populations, but as speakers of the same language, a large degree of cultural similarity can be assumed.

Not all variables in the *Atlas* are fit for Bayesian trait reconstruction. Many of them, such as the type of animal husbandry, display little or no variation across the subset of Indo-European peoples. Most variables are not binary, and only those that could be collapsed into a binary variable (for example, presence/absence of slavery) were used. Moreover, some traits were already investigated in previous studies: mode of marriage (Fortunato et al., 2006), the polygamy/monogamy part of family organization (Fortunato, 2011a), and marital residence (Fortunato, 2011b).

Following Fortunato (2011a,b) I used the Hittite language (or an Anatolian subfamily consisting of Hittite, Luvian and Lycian) as an outgroup for the Indo-European phylogenetic tree. I coded these as ambiguous for traits from the *Ethnographic Atlas*, but did not assign them any world religion as they predate all of those studied here.

Results

Marriage type

A maximum clade credibility (MCC) tree is the tree that, out of all those in a Bayesian sample, best summarizes it. Two MCC trees summarize the results of the replication of the Fortunato (2011a) study on marriage type (figure 1). The first tree and analysis use the same tree set as Fortunato (Pagel et al., 2007), and their results largely agree with her study, as expected. The root of the tree, representing the ancestral state of all Indo-European peoples, is reconstructed as monogamous with a posterior probability of 0.757 ± 0.180 (all results thus reported are the mean posterior probability across sample points, with standard deviation), similar to Fortunato's value of 0.70.

Quite different is the story told by the second tree, taken from the Bouckaert et al. (2012) tree set. In this reconstruction, the root has a 0.625 ± 0.379 probability of polygamy and conversely a 0.375 ± 0.379 probability of monogamy. There are also noticeable differences between the deep nodes of both trees, with those in the 2012 tree set carrying more trait state uncertainty. The discrepancy may be due to a difference in topology: in the 2012 set, the Balto-Slavic and Celtic-Germanic-Romance clades, which contain only monogamous societies, are seen as more likely to be sister groups, whereas the 2007 set puts the Indo-Iranian clade, in which most polygamous societies are found, as a sister group to the Balto-Slavic. However, the particular node serving as ancestor for these two groups has a low ($< 50\%$) occurrence in the sample, and the resulting topology may be an artifact of displaying only the MCC tree, which, although the most credible tree in the sample, may be quite different from the true topology.

I compared the likelihoods of the two models with the Bayes Factor (BF), which tests the relative strengths of two models by evaluating their likelihoods (Kass and Raftery, 1995; Fortunato, 2011a; Meade and Pagel, 2014). I used the harmonic mean of log-likelihood at the end of the chain as an estimate for the marginal likelihood. The harmonic mean of the 2007 chain is -13.493907, and that of the 2012 chain is -11.917916. Thus $\log(BF) = 2(-11.917916 - -13.493907) = 3.151982$, which is "positive" evidence that the 2012 trees constitute a more appropriate model for trait evolution of marriage type (Kass and Raftery, 1995). However, the harmonic mean may not be a very reliable estimation of the likelihood of a model, as it can and does vary greatly between independent runs.

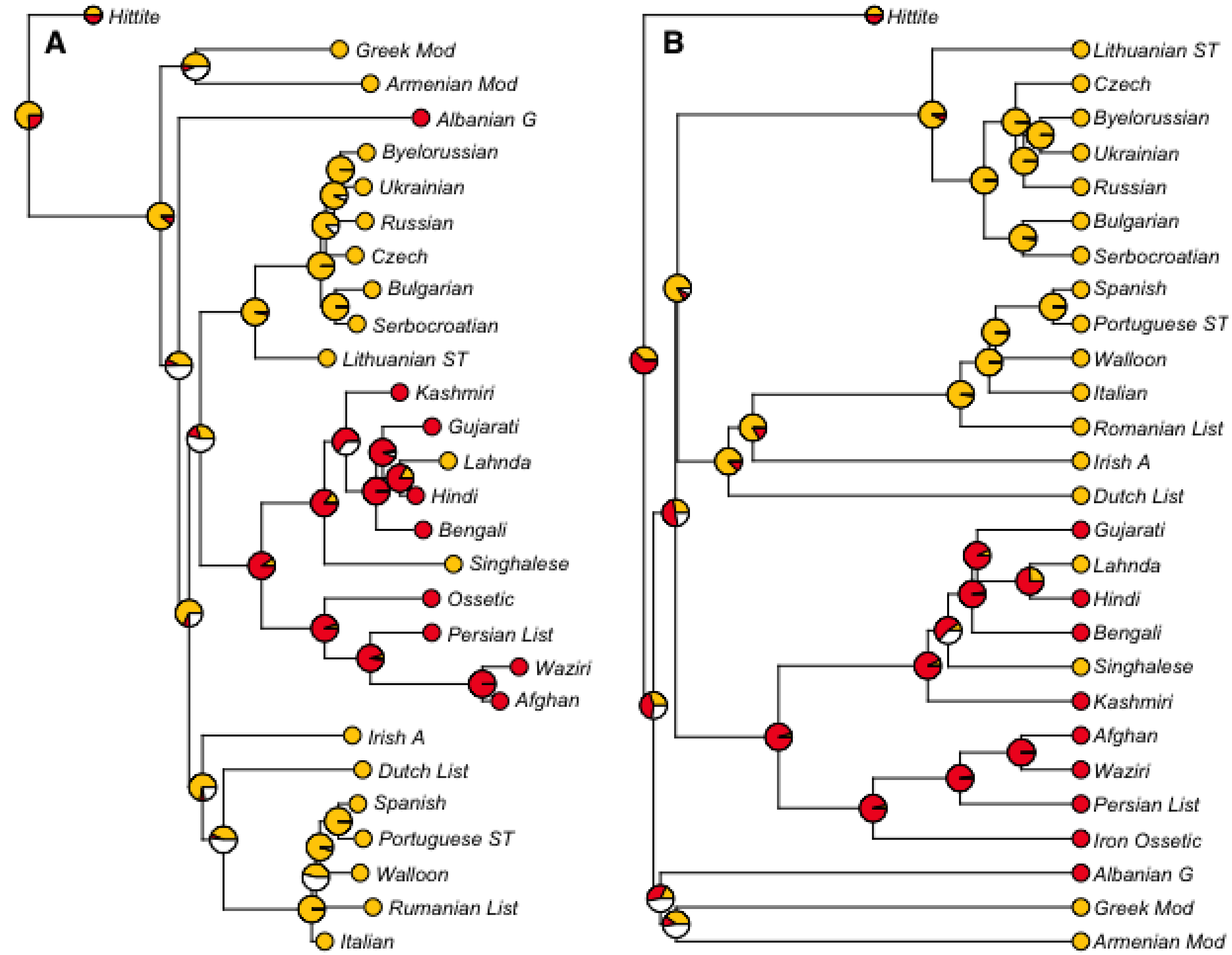


Figure 1: Reconstruction of ancestral marriage state in Indo-European cultures, using the Pagel et al. (2007) tree sample with branch lengths proportional to linguistic change (A), and the Bouckaert et al. (2012) tree sample with branch lengths proportional to time (B). The tips indicate the practice of monogamy (yellow) or polygyny (red). The internal nodes indicate the posterior probability of being in either state, with white indicating the proportion of sampled trees that did not include that node. The trees shown are the MCC trees obtained from their respective samples.

World religions

The results of the Christianity reconstruction on the 2012 set is shown in figure 2, with red bars indicating actual dates of conversion. The dates serve a comparison purpose and were not used in the reconstruction procedure. In principle, internal nodes that are reconstructed as Christian (blue) and that fall to the right of their associated red bars represent successful reconstruction, whereas those that fall to the left of the red bars, or that fall to the right and are reconstructed as non-Christian, represent failed reconstruction. As a caveat, it must be stressed that the red bars are but a very approximative attempt at dating a complex historical process.

Almost all Christian groups fall into a clade that includes Italic, Celtic, Germanic, Baltic and Slavic languages. In this clade, all recent societies are Christian, and only a few extinct languages are associated with non-Christian people (Old Prussian, Latin, Oscan, and Umbrian). These extinct languages are what makes the ancestors of the Baltic (consisting of Old Prussian, Lithuanian, and Latvian), Balto-Slavic (the Baltic plus the Slavic languages such as Polish and Slovenian), Romance (Latin and languages such as Spanish and Vlach) and Italic (Romance plus Oscan and Umbrian) reconstructed as probably non-Christian. By contrast, the ancestors of the Slavic, Germanic (Danish, English, etc.), and Celtic (Old Irish, Breton, etc.) groups are reconstructed as Christian with high probabilities, even though they fall further back in history than the times of Christianization of their descendent groups. In fact, three of these supposedly Christian ancestor nodes appear prior to the birth of Christianity itself. Thus, it seems that the presence of extinct languages, which can be considered in linguistics as analogous to the fossils of the biological world (Atkinson and Gray, 2005), have a great bearing on the results of the reconstruction. To test this, I ran a similar BayesTraits reconstruction analysis in which I pruned all extinct languages beforehand, except Hittite (figure 3). The loss of Latin and other non-Christian “fossils” caused the ancestors of the entire Balto-Slavic–Germanic–Celtic–Italic clade to be reconstructed as Christian with high probability, even though they are dated some 4000 years before Christ.

Outside of that clade, the only groups coded as unambiguously Christian in figure 2 are the Ossetic, Greek and Armenian (both modern and classical) ones. Besides the slightly too early Christian reconstruction of the ancestor of classical and modern Armenian, these languages are too isolated within their phylogenetic context to cause peculiar reconstructions. As for the root

Table 1: Number of times each model class was sampled by the RJMCMC procedure for the reconstruction of three world religions on a tree of Indo-European cultures. For each religion, the total number of sampled points is 99,000. q_{gain} and q_{loss} are the instantaneous rates of change towards the presence and absence, respectively, of the given religion.

	$q_{gain} = q_{loss} > 0$	$q_{gain} = 0$, $q_{loss} > 0$	$q_{gain} > 0$, $q_{loss} = 0$	$q_{gain} \neq q_{loss}$; $q_{gain}, q_{loss} > 0$
Christianity	35,286	3	63,372	339
Islam	67,784	3,989	26,903	324
Hinduism	7,225	88,433	5	3,337

and deep nodes of the tree, they are all, in an expected result, reconstructed as non-Christian with relatively high probability.

The RJMCMC approach distinguishes four models classes, as explained in the Materials and Methods section. How often model classes are sampled during an RJMCMC chain gives an indication of the most likely evolutionary scenarios. Table 1 gives the number of times each class was sampled during a chain run. In the case of Christianity, the most likely scenario according to BayesTraits is that gain of the religion is possible while loss is impossible (64% of the samples); the scenario in which both are possible and occur at the same rate appears in second position (36%). At almost no point during the run did the chain sample a scenario in which loss is possible, but gain is not.

Figure 4 displays the same MCC tree as used for the reconsruction of Islam. Islam is present in all recent Iranian-speaking societies except Ossetes; it also occurs in several Indic-speaking societies and in Albania. The clade consisting of the Iranian languages Kurdish, Baluchi, Persian, Tadzik, Afghan, and Waziri is fully Muslim, and as such their ancestors node are also reconstructed as Muslim, even though the earliest of these nodes fall much before the actual birth of Islam at the beginning of the 7th century. The most likely evolutionary model, according to the RJMCMC numbers, is the one in which gain and loss of religion are equal (68% of the samples; table 1).

The reconstruction of Hinduism (figure 5) showed a surprising result: even though Hindu societies are restricted to the Indic clade, the root of the entire Indo-European tree was reconstructed as Hindu with a 0.532 ± 0.250 probability. Dating the birth of the Hindu religion is difficult, as it is a synthesis of many Indian traditions; still, it is known that the ideas of modern Hinduism were codified over a period spanning from 300 BCE to 500

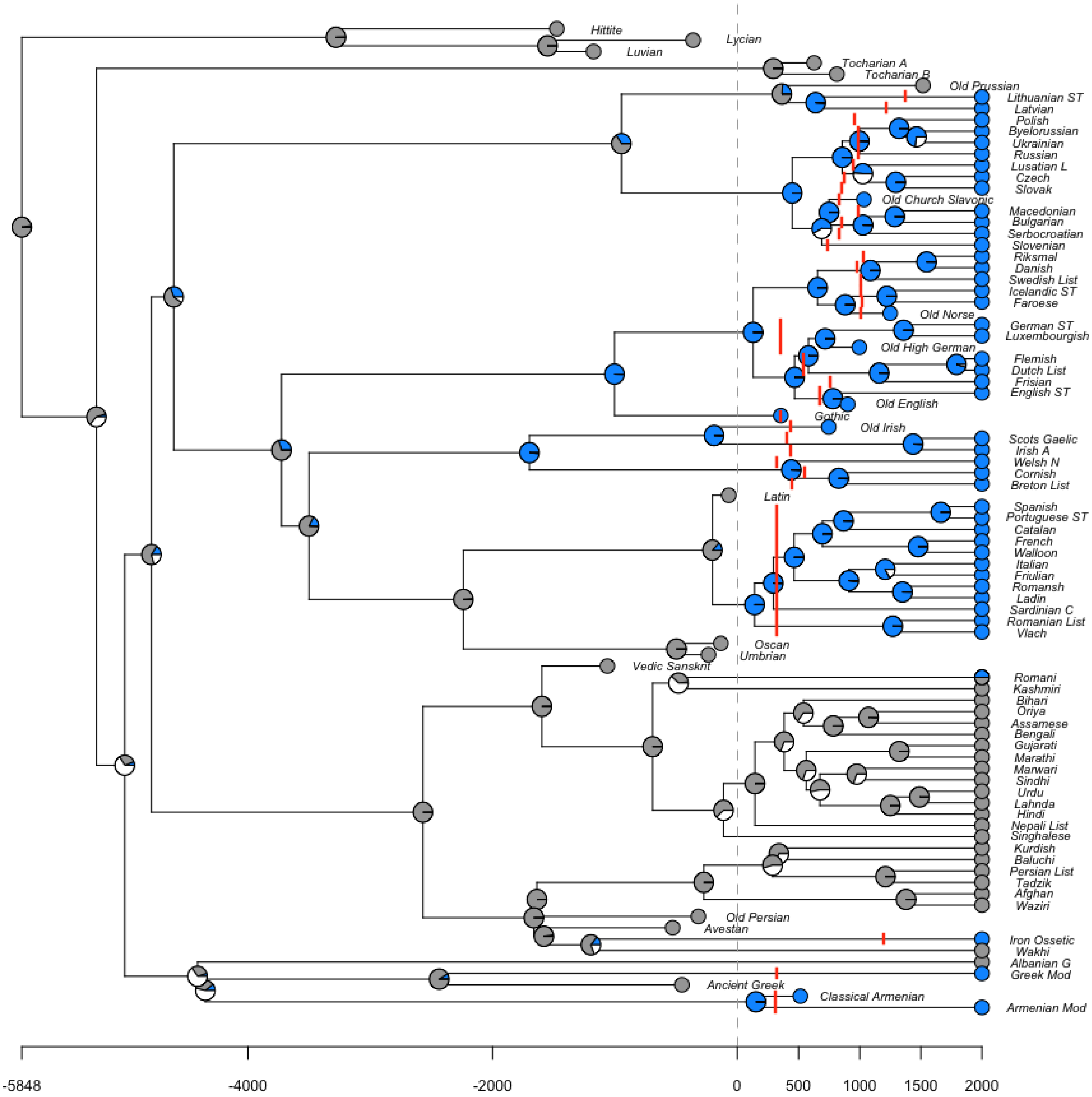

Hittite
Luvian
Lycian
Tocharian A
Tocharian B
Old Prussian
Lithuanian ST
Latvian
Polish
Byelorussian
Ukrainian
Russian
Lusatian L
Czech
Slovak
Old Church Slavonic
Macedonian
Bulgarian
Serbocroatian
Slovenian
Riksmal
Danish
Swedish List
Icelandic ST
Faroese
Old Norse
German ST
Luxembourgish
Old High German
Flemish
Dutch List
Frisian
English ST
Old English
Gothic
Old Irish
Scots Gaelic
Irish A
Welsh N
Cornish
Breton List
Latin
Spanish
Portuguese ST
Catalan
French
Walloon
Italian
Friulian
Romansh
Ladin
Sardinian C
Romanian List
Vlach
Oscan
Umbrian
Vedic Sanskrit
Romani
Kashmiri
Bihari
Oriya
Assamese
Bengali
Gujarati
Marathi
Marwari
Sindhi
Urdu
Lahnda
Hindi
Nepali List
Singhalese
Kurdish
Baluchi
Persian List
Tadzik
Afghan
Waziri
Old Persian
Avestan
Iron Ossetic
Wakhi
Albanian G
Greek Mod
Ancient Greek
Classical Armenian
Armenian Mod
-5848
-4000
-2000
0
500
1000
1500
2000

Figure 2: (*previous page*) Reconstruction of Christianity on a maximum clade credibility tree of Indo-European cultures. The tips indicate Christianity as a dominant religion (blue) or not (grey) or ambiguous (both). The internal nodes are pie charts representing the proportion of sampled states, during a BayesTraits RJMCMC analysis, in which the node was absent (white) or reconstructed as Christian or not. Approximate time in years is represented on the axis, with the assumption that the present is 2000 CE; the dashed line indicates the birth of Christianity at the beginning of the Common Era. Red bars indicate the approximate dates of Christianization for relevant groups; these bars were added posteriorly to the trait reconstruction.

CE (Lochtefeld, 2002). (Whether to code Vedic Sanskrit, dated circa 1000 BCE, as Hindu, is open to question, but both possibilities yield qualitatively similar reconstructions.) It is of course absurd that the common ancestor culture of all Indo-Europeans, thousands of years BCE, be Hindu. This result is ostensibly due to the most sampled model class (table 1), in which gain of religion is forbidden while loss is possible (89% of the samples). This forces the earliest nodes to be Hindu, as otherwise the religion could not appear at any point to create the pattern at the tips. Why this scenario was sampled so often may be due to the high number of loss events within the Indic clade (7 out of 9 cultures are coded as non-Hindu in this group).

I performed additional chain runs in which I fossilized, or forced, the root of the entire tree in the non-Hindu state. Incorporating this extra information gives results that are more in line with common sense, where the earliest node to be sampled as Hindu most of the time (posterior probability of 0.754 ± 0.208) is the ancestor of Sanskrit and all other Indic languages, though some older nodes are also given some probability of being in that state. However, the likelihood of the model with the fossilized node is lower: two independent runs of the chain with the fossil had harmonic means of -22.707044 and -23.134305, while two without it had harmonic means of -17.536721 and -18.084857. Averaging these values and computing the Bayes Factor gives a value of $\log(BF) = 10.21976$, "very strong" (Kass and Raftery, 1995) evidence that the model without the fossilized node is superior at explaining the distribution of trait data.

Other traits

The results of attempting reconstruction on trait data from the *Ethnographic Atlas* are of limited interest. In most cases, it is difficult or infeasible to assign

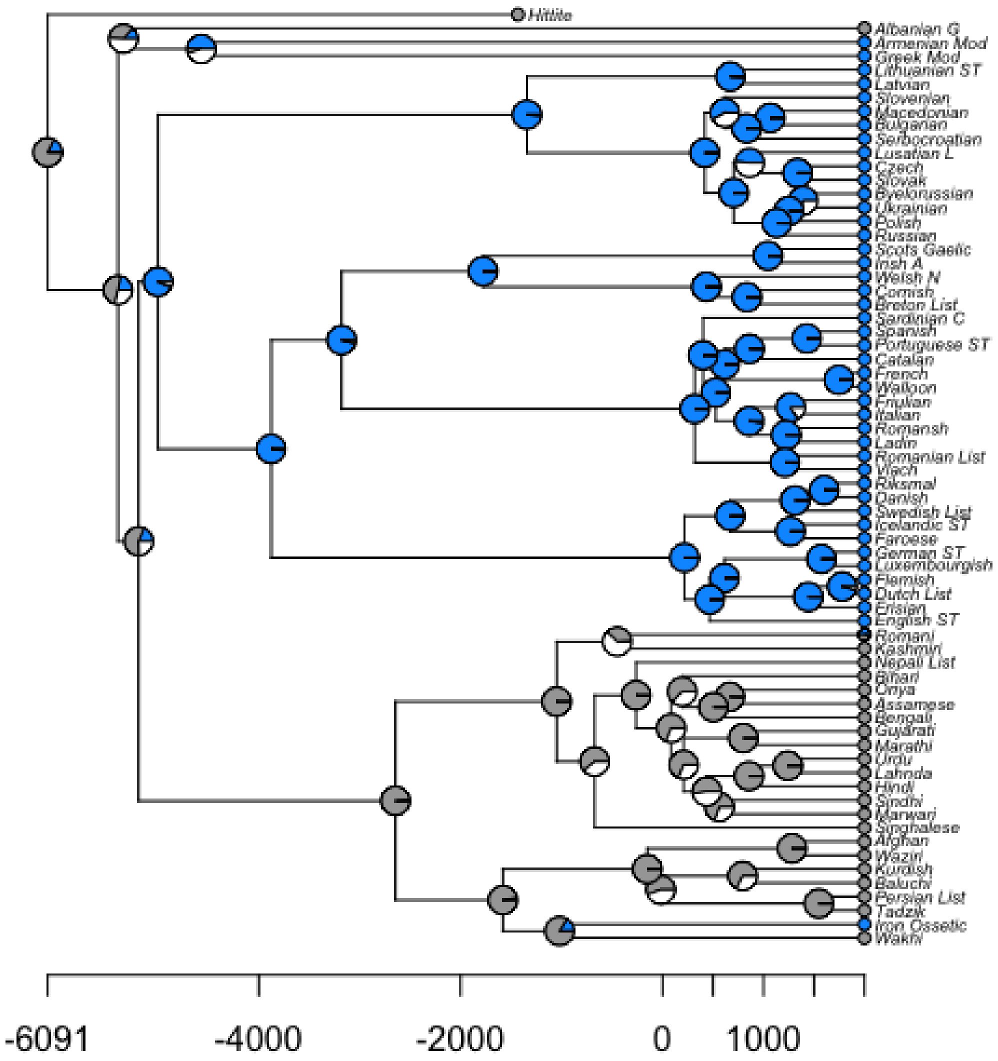


Figure 3: Results of a reconstruction analysis similar to that of figure 2, but in which extinct languages were removed beforehand.

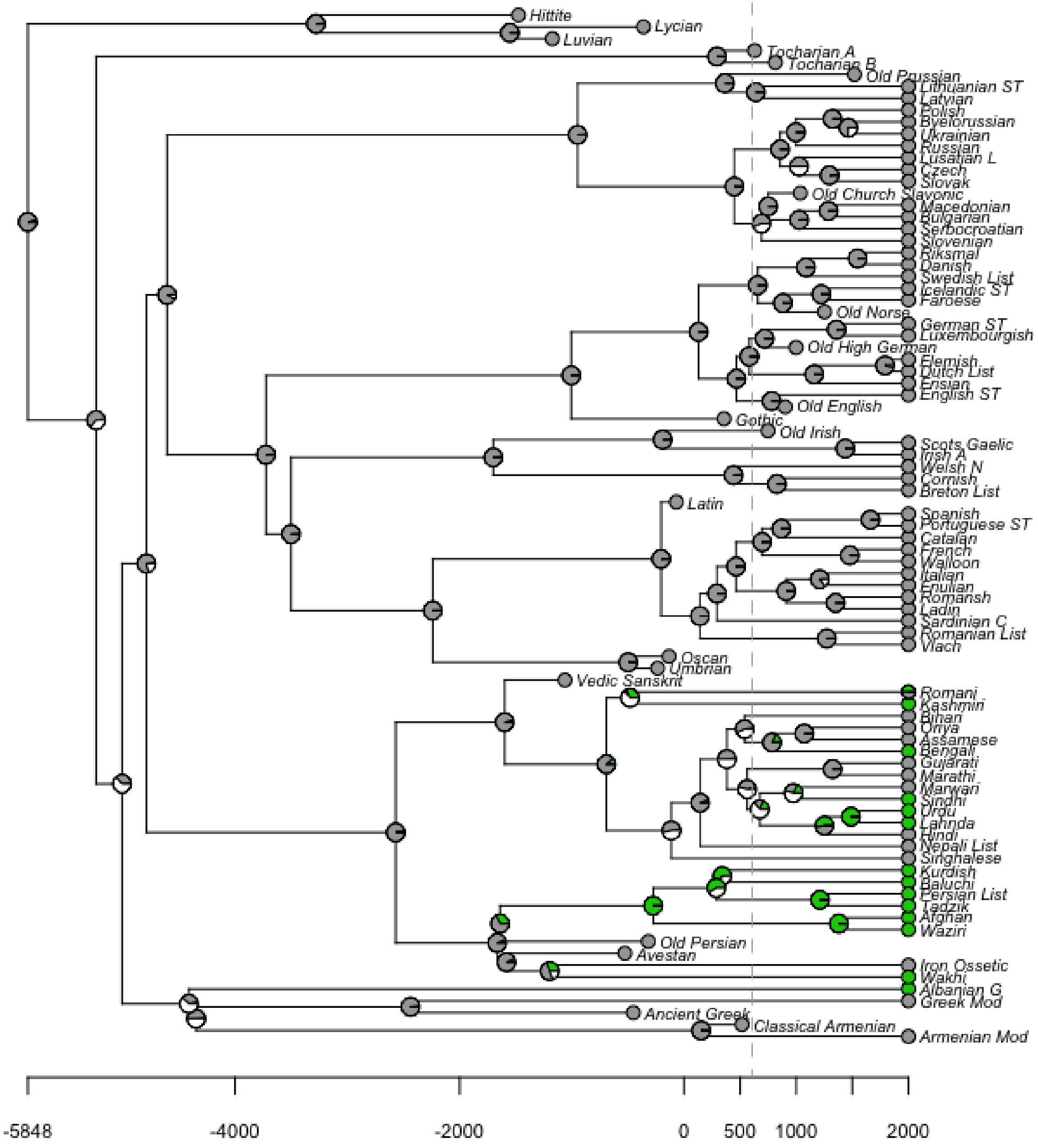


Figure 4: Reconstruction of Islam on a maximum clade credibility tree of Indo-European cultures. Green represents the presence of Islam, and grey its absence. The dashed line indicates the approximate birth of Islam circa 600. See figure 2 for a more detailed description.

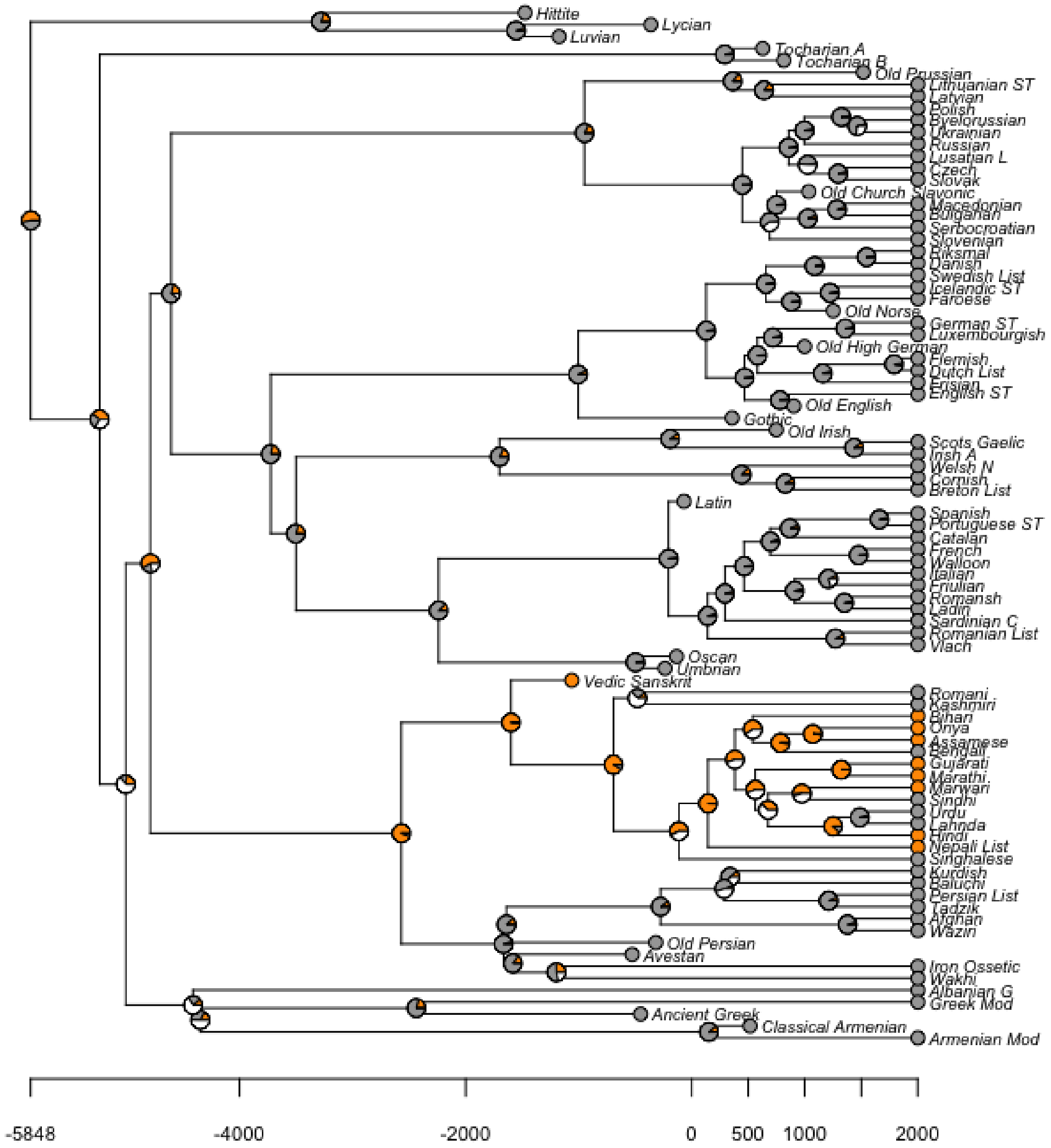


Figure 5: Reconstruction of Hinduism on a maximum clade credibility tree of Indo-European cultures. Orange represents the presence of Hinduism, and grey its absence. See figure 2 for a more detailed description.

dates to changes in trait state, disallowing proper benchmarking. Some of the MCC trees with trait reconstructions are shown in Appendix B. One case, slavery, is presented in figure 6. Slavery was coded in the *Atlas* as absent from most Indo-European societies except some Iranian ones as well as the ancient Latin- and Old Norse-speakers. Similarly to the case in which I excluded extinct language groups from the Christianity reconstruction, removing Latin and Old Norse alters the reconstruction of the deep nodes of the Germanic, Celtic and Italic clades. Without these languages, slavery is reconstructed to having probably disappeared from the ancestors of these clades much earlier than what historically happened.

Discussion

Phylogenetic comparative methods are an attractive alternative to conventional historical analysis for those researchers studying the evolutionary patterns of human culture. Certainly, they should not be rejected outright, but a tool is useful only insofar as it can provide reliable and consistent results. The goal, here, was to assess the reliability and consistency of Bayesian phylogenetic trait reconstruction in the Indo-European linguistic tree. The cultures represented by the languages in that phylogeny inhabit Europe and Western, Central and South Asia, areas that have a rich historical record. Thus the idea was not to gain additional knowledge on these peoples' history, but to study the method itself so that it can eventually be applied to areas of the world in which the historical record is poor.

In the introduction, I mentioned three important problems of Bayesian phylogenetic reconstruction: horizontal transmission, correctness of the tree, and incorporation of history. In light of the results, let us examine each of them in reverse order.

Incorporation of history

The cases of Christianity (figures 2-3) and slavery (figure 6), in which I removed extinct languages to find that the reconstruction substantially differed, emphasize the importance of including as much historical knowledge as possible in any given analysis. Extinct languages and societies such as the Latin-speaking Romans represent valuable “fossils” that provide information on traits that may have been lost in modern-day populations. Observing only

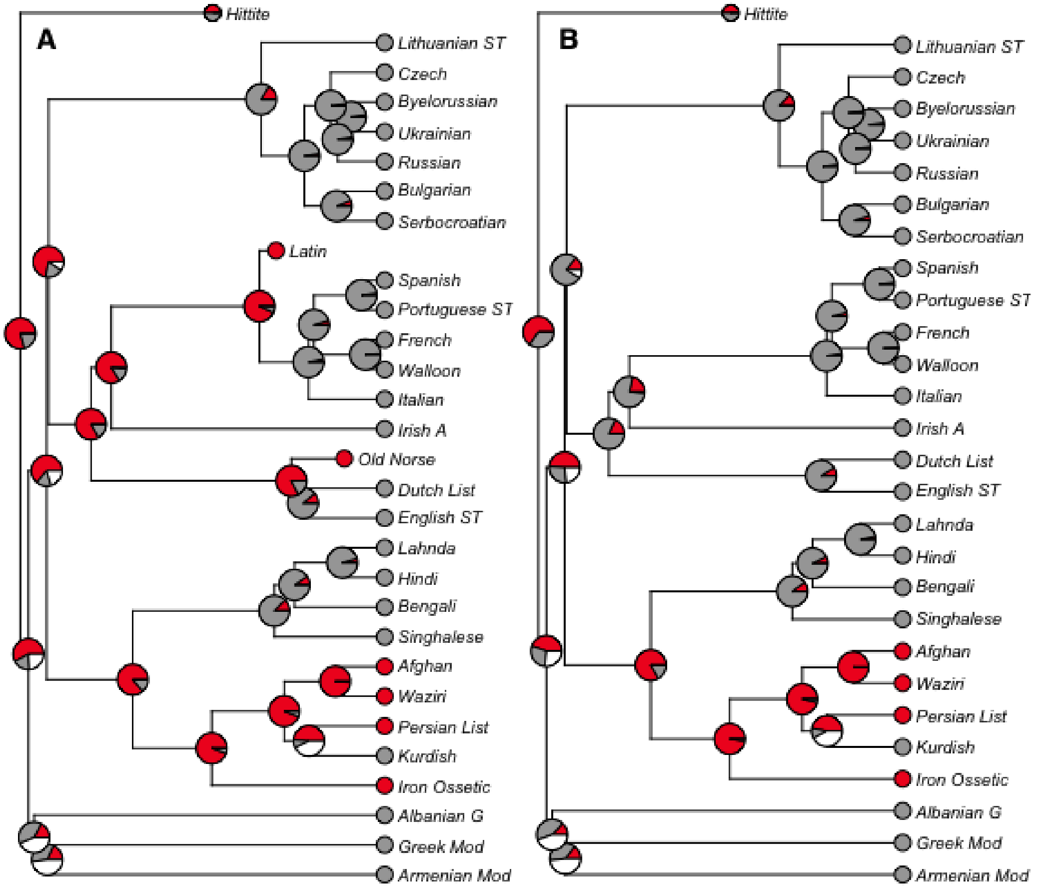


Figure 6: Reconstruction of slavery on two trees of Indo-European cultures, one including Latin- and Old Norse-speakers (A) and the other excluding them (B). Coding according to the *Ethnographic Atlas*. Red represents presence of slavery, grey its absence, and white the proportion of sampled states in which the node did not exist.

the current state, and ignoring history, could lead to false conclusions—such as that slavery has been absent from European societies ever since the split with the Indo-Iranian part of the group. This is of course absurd to anyone with basic historical knowledge, but it illustrates the ease with which trait reconstruction can be misleading, especially in the case of a trait or area whose past is poorly documented.

A similar conclusion can be made from the case where I used the "fossilizing" BayesTrait command to force the root of the tree to be non-Hindu. Fossilizing a node in BayesTrait is an explicit way to incorporate history and may thus improve the quality of a reconstruction. It will probably be obvious, however, that if one hopes to discover the trait state of an ancestor, one cannot fix it in a particular state beforehand. If phylogenetic trait reconstruction is to be used as an additional tool to study history, it cannot be required to only reach conclusions that have been reached by other historical methods. A good reconstruction analysis should aim to incorporate as much historical knowledge as possible in order to ask questions about ancestors that are truly unknown.

Phylogenetic trait reconstruction is probably not appropriate for the precise dating of historical events. In the Christianity (figure 2) and Islam (figure 4) cases, many nodes were reconstructed with the religion even though they fell earlier than actual times of conversion or even the birth of the religion. This is not a surprising result. Without detailed historical information, which would probably make a phylogenetic study unnecessary, high precision is unlikely. When a node is reconstructed, it should be viewed more as information about an ancient people, such as the proto-Indo-Europeans, than about a time period. The method is perhaps more applicable to the study of abstract evolutionary patterns such as the coevolution of traits (e.g. Watts et al., 2015, 2016) or the dating of events over broad time scales such as the origin of the Indo-European group (Bouckaert et al., 2012).

Correctness of the tree

Bayesian methods have the advantage of making phylogenetic uncertainty explicit by working with a tree set rather than a single, possibly erroneaous tree (Pagel et al., 2004; Mace and Holden, 2005). This uncertainty is illustrated in the figures by the white portion of reconstructed nodes, indicating the proportion of samples in the MCMC chain where the node did not exist. The resulting probability that a node exists in the true phylogeny therefore puts an upper bound on reconstruction probabilities (Fortunato, 2011a). This is a feature of Bayesian methods, not a problem, although it means that tree sets with a lot of uncertain nodes may be limited in their use to reconstruction studies.

What is of greater concern is the correctness of the tree set itself as compared to other sets made from different data or with different computational tools. The marriage example (figure 1) highlights the fact that using two different, albeit similar, tree sets can lead to strikingly contrasting results. It may be wise, when attempting to reconstruct ancestral states, to perform the analysis on a number of tree sets, including, when possible, some built from linguistic as well as genetic data.

Horizontal transmission

While the horizontal transmission of cultural traits does not necessarily invalidate phylogenetic studies of cultural evolution (Collard et al., 2006; Greenhill et al., 2009), it certainly is a prevalent phenomenon and must be taken into account. World religions are a prime example of a trait that is often transmitted independently of ancestry, by processes such as evangelization and military conquest, and this may impair the accuracy of a phylogenetic reconstruction. On the other hand, religions are also transmitted vertically as people generally adopt the belief of their immediate ancestors. In figure 2, many shallow nodes are correctly reconstructed as Christian, particularly in the Romance clade. After the Roman Empire became a Christian state, the new states that emerged from its collapse kept that religion, unless believers of another religion (Islam) converted them.

Of course, world religions, as very specific cultural practices, can in fact not occur as independent mutations—in the way that more general traits like high gods or slavery can—but only as a result of horizontal transfer. The presence of Christianity in Armenia, Greece or Ossetia, although seemingly independent from the other groups in the phylogeny, certainly did not happen in isolation. Yet, in an abstract sense, the appearance of any trait can be likened to a “mutation” event happening along the cultural lineage of a group. Horizontal transmission merely provides an additional mutagenesis mechanism. It is not independent of cultural similarity and geography, and this may help fitting the resulting trait distribution on a linguistic tree, since linguistic evolution also captures the influence of these factors (Cavalli-Sforza, 1997).

Cultural traits may be more or less prone to horizontal transmission, much like, in linguistics, certain words may be more readily borrowed than others (Pagel et al., 2007). Many are also associated with the spread of overarching cultural practices like religion. In Testart (2013)'s critique of the Fortunato et al. (2006) study on marriage transfer, he observes that the differences in marriage transfer practices match almost exactly those in religion (almost all Christian and all Hindu populations have dowry, whereas Muslim populations have bridewealth). The adoption of monogamy could be correlated with that of Christianity: all Christian cultures in figure 1 are monogamous, and almost all monogamous cultures are Christian. The practice of circumcision (Appendix B) seems to match the distribution of Islam (figure 4).

Researchers must be aware of the degree to which a trait may be horizontally transferable when studying them in a phylogenetic framework. Methods allowing some degree of reticulation, which are being developed in biology, also represent an avenue that should be explored by cultural evolutionists (Gray et al., 2007; Lewens, 2013).

Conclusion

According to Cooper et al. (2016), phylogenetic methods carry a risk of being used by researchers who are not sufficiently aware of their issues, biases and limitations. While this is true of any domain of application, it is perhaps even more so in cultural evolution. The findings presented in this report stress the importance of thinking carefully about the problems inherent to attempting the reconstruction of cultural traits on phylogenetic trees. This way, more confidence can be placed in the use of what can certainly be a valuable tool.

Bibliography

Atkinson QD, Gray RD. 2005. Curious parallels and curious connections—Phylogenetic thinking in biology and historical linguistics. *Systematic Biology* 54: 513-526.

Berend N (ed.). 2007. *Christianization and the Rise of Christian Monarchy: Scandinavia, Central Europe and Rus' c. 900-1200.* Cambridge University Press.

Bouckaert R, Lemey P, Dunn M, Greenhill SJ, Alekseyenko AV, Drummond AJ, Gray RD, Suchard MA, Atkinson QD. 2012. Mapping the origins and expansion of the Indo-European language family. *Science* 337: 957-960.

Cavalli-Sforza LL. 1997. Genes, peoples, and languages. *PNAS* 94: 7719-7724.

Collard M, Shennan SJ, Tehrani JJ. 2006. Branching, blending, and the evolution of cultural similarities and differences among human populations. *Evolution and Human Behavior* 27: 169-184.

Cooper N, Thomas GH, FitzJohn RG. 2016. Shedding light on the 'dark side' of phylogenetic comparative methods. *Methods in Ecology and Evolution* 7: 693-699.

Currie TE, Greenhill SJ, Gray RD, Hasegawa T, Mace R. 2010. Rise and fall of political complexity in island South-East Asia and the Pacific. *Nature* 467: 801-804.

Dyen I, Kruskal JB, Black P. 1992. An Indoeuropean classification: A lexicostatistical experiment. *Transactions of the American Philosophical Society* 82.

Fortunato L. 2011a. Reconstructing the history of marriage strategies in Indo-European–speaking societies: monogamy and polygyny. *Human Biology* 83: 87-105.

Fortunato L. 2011b. Reconstructing the history of residence strategies in Indo-European–speaking societies: neo-, uxori-, and virilocality. *Human Biology* 83: 107-128.

Fortunato L, Holden C, Mace R. 2006. From bridewealth to dowry? A Bayesian estimation of ancestral states of marriage transfers in Indo-European groups. *Human Nature* 17: 355-376.

Gould SJ. 1987. *An urchin in the storm.* New York: Norton.

Gray JP. 1999. A corrected *Ethnographic Atlas. World Cultures* 10: 24-85.

Gray RD, Greenhill SJ, Ross RM. 2007. The pleasures and perils of Darwinizing culture (with phylogenies). *Biological Theory*, 2(4).

Greenhill SJ, Currie TE, Gray RD. 2009. Does horizontal transmission invalidate cultural phylogenies? *Proceedings of the Royal Society B* 276: 2299-2306.

Huelsenbeck JP, Ronquist F, Nielsen R, Bollpack JP. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294: 2310-2314.

Jordan F, Gray RD, Greenhill SJ, Mace R. 2009. Matrilocal residence is ancestral in Austronesian societies. *Proceedings: Biological Sciences* 276: 1957-1964.

Kass RE, Raftery AE. 1995. Bayes Factors. *Journal of the American Statistical Association* 90: 773-795.

Kurian GT (ed.). 2011. *Encyclopedia of Christian Civilization.* Wiley-Blackwell.

Lewens T. 2013. Cultural Evolution. In: *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Edward N. Zalta (ed.), `plato.stanford.edu/archives/spr2013/entries/evolution-cultural/`.

Lewis GL. 1991. Čingāne. *The Encyclopaedia of Islam* (2nd ed.). Brill, p. 40-41.

Lochtefeld JG. 2002. *The Illustrated Enclyclopedia of Hinduism.* Rosen.

Mace R, Holden CJ. 2005. A phylogenetic approach to cultural evolution. *Trends in Ecology and Evolution* 20: 116-121.

Mace R, Pagel M. 1994. The comparative method in anthropology. *Current Anthropology* 35: 549-564.

Martin RC (ed. in chief). 2016. *Encyclopedia of Islam and the Muslim World.* Macmillan Reference USA.

Meade A, Pagel M. 2014. BayesTraits V2 Manual. Available from `www.evolution.rdg.ac.uk/BayesTraitsV2.0Files/TraitsV2Manual.pdf`. Retrieved 19 July 2016.

Murdock GP. 1967. Ethnographic Atlas: A summary. *Ethnology* 6: 109-236.

Murdock GP, White DR. 1969. Standard cross-cultural sample. *Ethnology* 9: 329-369.

Opie C, Shultz S, Atkinson QD, Currie T, Mace R. 2014. Phylogenetic reconstruction of Bantu kinship challenges Main Sequence Theory of human social evolution. *PNAS* 111: 17414-17419.

Pagel M, Atkinson QD, Calude AS, Meade A. 2013. Ultraconserved words point to deep language ancestry across Eurasia. *PNAS* 110: 8471-8476.

Pagel M, Atkinson QD, Meade A. 2007. Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* 449: 717-721.

Pagel M, Meade A. 2006. Bayesian analysis of correlated evolution of discrete characters by revesible-jump Markov chain Monte Carlo. *The American Naturalist* 167: 808-825.

Pagel M, Meade A, Barker D. 2004. Bayesian estimation of ancestral character states on phylogenies. *Systems Biology* 53: 673-684.

Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289-290.

Peoples HC, Duda P, Marlowe FW. 2016. Hunter-gatherers and the origin of religion. *Human Nature* doi:10.1007/s12110-016-9260-0.

R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Ringe DA. 1995. 'Nostratic' and the factor of change. *Diachronica* 12: 53-74.

Ringe D, Warnow T, Taylor A. 2002. Indo-European and computational cladistics. *Transactions of the Philological Society* 100: 59-129.

Schliep KP. 2011. phangorn: phylogenetic analysis in R. *Bioinformatics* 27(4): 592-593.

Tëmkin I, Eldredge N. 2007. Phylogenetics and material culture evolution. *Current Anthropology* 48: 146-153.

Testart A. 2013. Reconstructing social and cultural evolution: The case of dowry in the Indo-European area. *Current Anthropology* 54: 23-50.

Waldman C, Mason C. 2006. *Encyclopedia of European Peoples.* Infobase Publishing.

Walker RS, Hill KR, Flinn MV, Ellsworth RM. 2011. Evolutionary history of hunter-gatherer marriage practices. *PLoS ONE* 6: e19066.

Walker RS, Wichmann S, Mailund T, Atkisson CJ. 2012. Cultural phylogenetics of the Tupi language family in lowland South America. *PLoS ONE* 7: e35025.

Watts J, Greenhill SJ, Atkinson QD, Currie TE, Bulbulia J, Gray RD. 2015. Broad supernatural punishment but not moralizing high gods precede the evolution of political complexity in Austronesia. *Proceedings of the Royal Society B* 282: 20142556.

Watts J, Sheehan O, Atkinson QD, Bulbulia J, Gray RD. 2016. Ritual human sacrifice promoted and sustained the evolution of stratified societies. *Nature* 532: 228-231.

Appendix A: Table of dates of Christianization

Language group	**Date**	**Source**
Armenian Mod	301	Kurian (2011, p. 120)
Breton List	c. 450	Waldman and Mason (2006, p. 80)
Bulgarian	865	Waldman and Mason (2006, p. 103)
Byelorussian	988	Berend (2007); Waldman and Mason (2006, p. 669)
Catalan	313	Waldman and Mason (2006, p. 658)
Classical Armenian	301	Kurian (2011, p. 120)
Cornish	6th c.	Waldman and Mason (2006, p. 181)
Czech	884	Berend (2007); Waldman and Mason (2006, p. 200)
Danish	963	Berend (2007); Waldman and Mason (2006, p. 840)
Dutch List	6th c.	Waldman and Mason (2006, p. 327)
English ST	680s	Waldman and Mason (2006, p. 236)
Faroese	1000	Berend (2007, p. 35-36)
Flemish	6th c.	Waldman and Mason (2006, p. 327)
French	313	Waldman and Mason (2006, p. 658)
Frisian	8th c.	Waldman and Mason (2006, p. 283)
Friulian	313	Waldman and Mason (2006, p. 658)
German ST	341	Waldman and Mason (2006, p. 327-340)
Gothic	341	Waldman and Mason (2006, p. 327-340)
Greek Mod	313	Waldman and Mason (2006, p. 658)
Icelandic ST	1000	Berend (2007, p. 35-36)
Irish A	431	Waldman and Mason (2006, p. 442)
Iron Ossetic	12th-13th c.	Waldman and Mason (2006, p. 572)
Italian	313	Waldman and Mason (2006, p. 658)
Ladin	313	Waldman and Mason (2006, p. 658)
Latvian	1214	Waldman and Mason (2006, p. 54)
Lithuanian ST	1386	Waldman and Mason (2006, p. 489)
Lusatian L	958	Waldman and Mason (2006, p. 778-779)

Luxembourgish	341	Waldman and Mason (2006, p. 327-340)
Macedonian	late 10th c.	Waldman and Mason (2006, p. 767)
Old Church Slavonic	831	Berend (2007)
Old English	680s	Waldman and Mason (2006, p. 236)
Old High German	341	Waldman and Mason (2006, p. 327-340)
Old Irish	431	Waldman and Mason (2006, p. 442)
Old Norse	1000	Berend (2007, p. 35-36)
Polish	966	Waldman and Mason (2006, p. 599)
Portuguese ST	313	Waldman and Mason (2006, p. 658)
Riksmal	1000-1020	Berend (2007); Waldman and Mason (2006, p. 840)
Romanian List	313	Waldman and Mason (2006, p. 767)
Romansh	313	Waldman and Mason (2006, p. 767)
Russian	late 10th c.	Waldman and Mason (2006, p. 767)
Sardinian C	313	Waldman and Mason (2006, p. 767)
Scots Gaelic	397	Waldman and Mason (2006, p. 718)
Serbocroatian	9th c.	Waldman and Mason (2006, p. 729)
Slovak	863	Waldman and Mason (2006, p. 768)
Slovenian	c. 740	Waldman and Mason (2006, p. 774)
Spanish	313	Waldman and Mason (2006, p. 658)
Swedish List	1000	Berend (2007)
Ukrainian	988	Berend (2007); Waldman and Mason (2006, p. 669)
Vlach	313	Waldman and Mason (2006, p. 658)
Walloon	313	Waldman and Mason (2006, p. 658)
Welsh N	313	Waldman and Mason (2006, p. 872)

Appendix B: Reconstruction of some traits from the *Ethnographic Atlas*

MCC trees displaying the results of Bayesian reconstruction for six traits from the *Ethnographic Atlas* are shown on the next page. In each case, the portion of an internal node that is white signifies the proportion of samples in the MCMC chain in which the node was absent. Colored tips and portions of internal nodes represent a "yes" answer to a binary question about the trait, and grey tips and portions of internal nodes represent a "no" answer. The trait questions are as follows:

A) Are caste distinctions present in this culture?

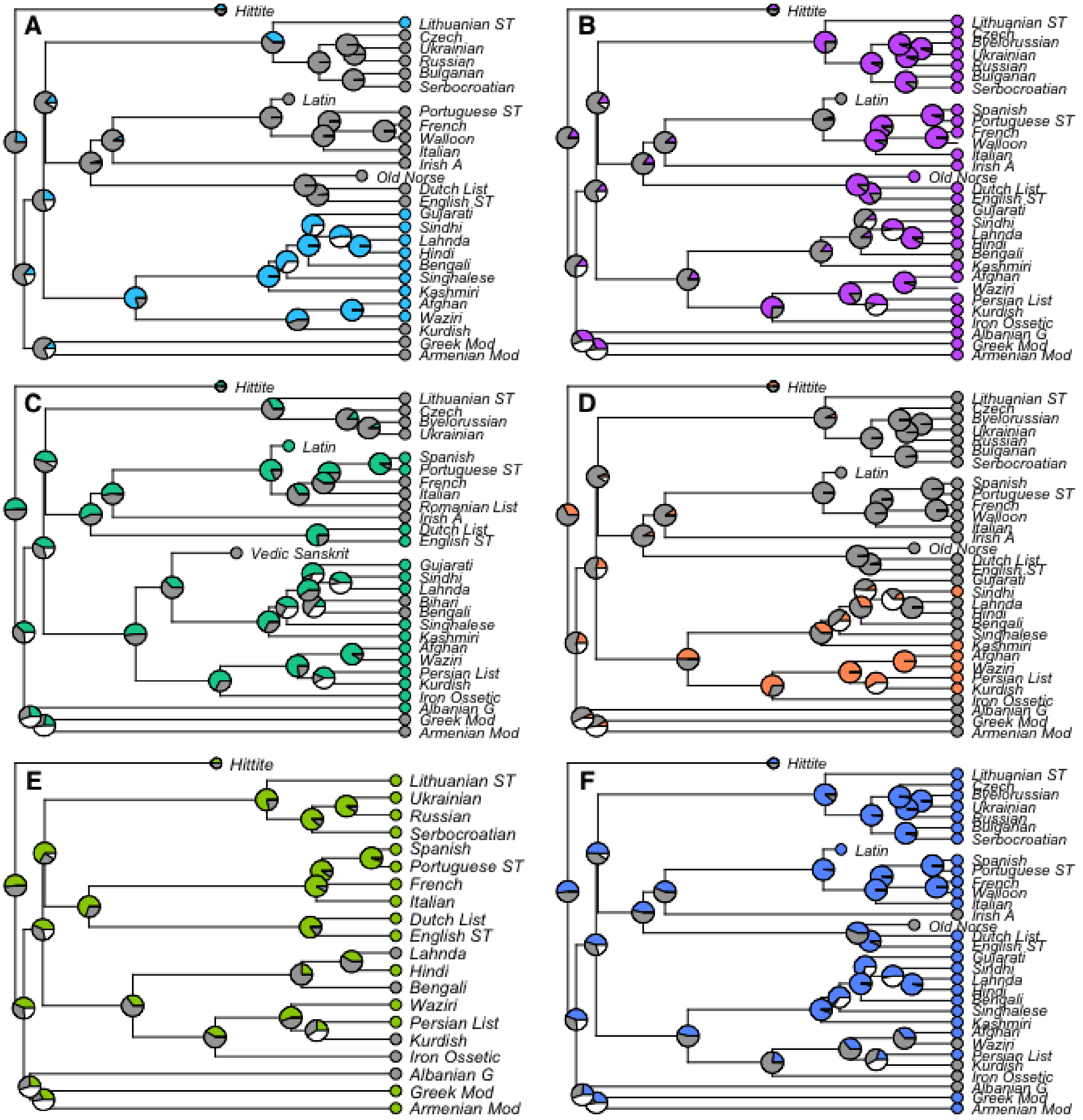
B) Is a high god, i.e. a divinity believed to be the creator or governor of reality, present?

C) Is marriage between first cousins acceptable?

D) Is male genital mutilation (circumcision) practiced?

E) Are local leaders chosen in a nonhereditary process (democratic, or by appointment)?

F) Does the culture form a politically complex large state?


A
Hittite
Lithuanian ST
Czech
Ukrainian
Russian
Bulgarian
Serbocroatian
Latin
Portuguese ST
French
Walloon
Italian
Irish A
Old Norse
Dutch List
English ST
Gujarati
Sindhi
Lahnda
Hindi
Bengali
Singhalese
Kashmiri
Afghan
Waziri
Kurdish
Greek Mod
Armenian Mod
B
Hittite
Lithuanian ST
Czech
Byelorussian
Ukrainian
Russian
Bulgarian
Serbocroatian
Latin
Spanish
Portuguese ST
French
Walloon
Italian
Irish A
Old Norse
Dutch List
English ST
Gujarati
Sindhi
Lahnda
Hindi
Bengali
Kashmiri
Afghan
Waziri
Persian List
Kurdish
Iron Ossetic
Albanian G
Greek Mod
Armenian Mod
C
Hittite
Lithuanian ST
Czech
Byelorussian
Ukrainian
Latin
Spanish
Portuguese ST
French
Italian
Romanian List
Irish A
Dutch List
English ST
Vedic Sanskrit
Gujarati
Sindhi
Lahnda
Bihari
Bengali
Singhalese
Kashmiri
Afghan
Waziri
Persian List
Kurdish
Iron Ossetic
Albanian G
Greek Mod
Armenian Mod
D
Hittite
Lithuanian ST
Czech
Byelorussian
Ukrainian
Russian
Bulgarian
Serbocroatian
Latin
Spanish
Portuguese ST
French
Walloon
Italian
Irish A
Old Norse
Dutch List
English ST
Gujarati
Sindhi
Lahnda
Hindi
Bengali
Singhalese
Kashmiri
Afghan
Waziri
Persian List
Kurdish
Iron Ossetic
Albanian G
Greek Mod
Armenian Mod
E
Hittite
Lithuanian ST
Ukrainian
Russian
Serbocroatian
Spanish
Portuguese ST
French
Italian
Dutch List
English ST
Lahnda
Hindi
Bengali
Waziri
Persian List
Kurdish
Iron Ossetic
Albanian G
Greek Mod
Armenian Mod
F
Hittite
Lithuanian ST
Czech
Byelorussian
Ukrainian
Russian
Bulgarian
Serbocroatian
Latin
Spanish
Portuguese ST
French
Walloon
Italian
Irish A
Old Norse
Dutch List
English ST
Gujarati
Sindhi
Lahnda
Hindi
Bengali
Singhalese
Kashmiri
Afghan
Waziri
Persian List
Kurdish
Iron Ossetic
Albanian G
Greek Mod
Armenian Mod