Does horizontal transmission invalidate cultural phylogenies?

Simon J. Greenhill^{1,*}, Thomas E. Currie² and Russell D. Gray¹

¹Department of Psychology, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

²Department of Anthropology, University College London, 14 Tavistock Street, London WC1H 0BW, UK

Phylogenetic methods have recently been applied to studies of cultural evolution. However, it has been claimed that the large amount of horizontal transmission that sometimes occurs between cultural groups invalidates the use of these methods. Here, we use a natural model of linguistic evolution to simulate borrowing between languages. The results show that tree topologies constructed with Bayesian phylogenetic methods are robust to realistic levels of borrowing. Inferences about divergence dates are slightly less robust and show a tendency to underestimate dates. Our results demonstrate that realistic levels of reticulation between cultures do not invalidate a phylogenetic approach to cultural and linguistic evolution.

Keywords: cultural phylogenetics; horizontal transfer; borrowing; simulation

1. INTRODUCTION

The only figure in Darwin's (1859) *Origin of species* is a tree. This figure represents Darwin's view of evolution as a process of descent with modification from a common ancestor. Since the publication of that book, there has been an ongoing debate about how evolutionary ideas can be applied to cultural and linguistic changes (Aunger 2000; Laland & Brown 2002; Carneiro 2003). The last few decades in evolutionary biology have seen the realization that 'tree thinking' (O'Hara 1988) is not just a way of describing evolution, but a way of testing evolutionary hypotheses. Evolutionary biologists have developed a powerful set of methods, phylogenetics, to answer evolutionary questions (Harvey & Pagel 1991; Huelsenbeck & Rannala 1997; Pagel 1999). The potential of phylogenetic methods for evolutionary analyses of culture has not gone unnoticed, with a recent proliferation of studies exploring everything from population movements (Gray & Jordan 2000; Holden 2002; Gray & Atkinson 2003; Rexová *et al.* 2003, 2006; Greenhill & Gray 2005; Holden & Gray 2006; Gray *et al.* 2009) to the evolution of material culture (Tehrani & Collard 2002; Jordan & Shennan 2003; Darwent & O'Brien 2006; Harmon *et al.* 2006; Neff 2006; Tëmkin & Eldredge 2007; Coward *et al.* 2008), social organization (Holden & Mace 2005; Mace & Jordan 2005; Fortunato *et al.* 2006; Moylan *et al.* 2006; Jordan *et al.* in press) and broader questions about language evolution (Dunn *et al.* 2005; Pagel *et al.* 2007; Atkinson *et al.* 2008).

The justification for cultural phylogenies is twofold. First, there are many similarities between biological and cultural evolution (Mace & Pagel 1994; Atkinson & Gray 2005; Mace & Holden 2005). Second, many of the questions asked by anthropologists are similar to those asked by evolutionary biologists, and thus can be answered using the robust statistical and inferential framework provided by phylogenetic methods (Gray *et al.* 2007).

However, like other attempts to Darwinize culture, the phylogenetic approach has been highly controversial (Gould 1987, 1991; Bateman *et al.* 1990; Moore 1994; Bellwood 1996; Borgerhoff Mulder 2001; Holden & Shennan 2005; Tëmkin & Eldredge 2007). In particular, it is frequently argued that the horizontal transmission of traits is rampant in cultural evolution:

Human cultural evolution proceeds along paths outstandingly different from the ways of genetic change... Trees are correct topologies of biological evolution... In human cultural evolution, on the other hand, transmission and anastomosis are rampant. Five minutes with a wheel, a snowshoe, a bobbin, or a bow and arrow may allow an artisan of one culture to capture a major achievement of another.

(Gould 1987, p. 70)

According to the critics, the borrowing of traits between cultures violates the assumptions of the phylogenetic method, and may thus lead to inaccurate results or even completely invalidate the approach. Instead of phylogenetic analyses of cultures, these critics argue for a 'rhizotic' view of culture where each society derives from multiple parents (Terrell 1988; Moore 1994; Terrell *et al.* 2001).

This highly polarized debate about the problems caused by cultural reticulation will not be settled by further hand-waving and armchair speculation. The central issue is not whether phylogenetics is appropriate on *a priori* grounds, but how robust phylogenetic inferences are to horizontal transmission. The specific questions that need to be answered are (i) what effect do different levels of horizontal transmission have on the accuracy of phylogenetic estimates and (ii) are the problematic levels of horizontal transmission common in real situations? In this study, we will answer these questions by using computer simulations to explore the impact horizontal transmission has on our ability to infer phylogenetic tree topologies and divergence dates accurately.

* Author for correspondence (s.greenhill@auckland.ac.nz).

Figure 1. The ‘true’ tree topologies used to synthesize the data, showing branch lengths and calibration points. (a) Tree 1 has a root age of 5111.14 years and (b) tree 2 has a root age of 5013.12 years.

2. MATERIAL AND METHODS

To investigate the impact of horizontal transmission on phylogenetic estimates, we simulated the evolution of languages under a natural model of language evolution—the stochastic Dollo model (SDM; Nicholls & Gray 2006, 2008). Simulations are a standard method for analysing the sensitivity of phylogenetic methods under a set of assumptions (Penny *et al.* 1992; Huelsenbeck 1995). First, we generated data by simulating the evolution of linguistic traits on two different tree topologies under varying degrees of horizontal transmission. Then, we evaluated how horizontal transmission affects the ability of these methods that do not account for horizontal transmission to recover the ‘true’ tree topology. Finally, we explored how borrowing affects inferences taken from the tree structure by attempting to estimate the age at the root of the trees, and how this varies from the root age of the true trees.

(a) Input trees

To act as our true histories, we constructed two topologically dissimilar trees from the existing language datasets (figure 1). The shape of a tree reflects the overall evolutionary process (Moore & Heard 1997). It contains information about population history, including the pattern and rates of speciation and the amount of change between populations. The effect of borrowing is likely to depend on this shape. The first tree has a ‘balanced’ topology where most nodes have an equal number of daughter cultures and relatively long internal branches. By contrast, the second topology is less balanced with a chained pattern of descent and much shorter internal branches.

(b) Data simulation

Data were simulated on both topologies using the SDM of lexical evolution, implemented in the software package TRAITLAB (Nicholls & Gray 2006, 2008). This models the

birth and death of traits along each lineage, as well as simulating the borrowing of traits between lineages (figure 2). A dataset is simulated under the SDM by starting with a language (comprising a set of traits) at the root of the tree. Languages change as they move along the lineages of the tree towards the tree tips by having new traits born and existing traits die. Traits are born at a birth rate, λ , and die at a death rate, v . In our simulations, λ and v were kept constant throughout. The value of v was chosen so that, along any lineage, a mean proportion of 0.2 traits will be lost per 1000 years. This proportion approximates the estimated rate at which trait classes from the 200-item basic-vocabulary ‘Swadesh’ word lists are lost (Swadesh 1952; Embleton 1986; Pagel & Meade 2006). At each branching point or node in the input tree, the trait sets entering the split are copied identically into both daughter languages. The traits do not interact with each other, and the same trait cannot be born independently into two or more lineages. Thus, if a trait is found in two separate languages, then it has either been inherited from a common ancestral language or been borrowed from one language into another. This is a realistic assumption as it is rare that two languages will independently invent the same word with the same meaning. At the end of the simulation, a data matrix of ones and zeros is produced, and can be thought of as representing the presence or absence of words in cognate sets for each extant language.

In the simulations, traits could also be borrowed from one language into another, at each instant, according to a specified rate. When a trait is borrowed from one language into another, an identical copy of the selected trait is duplicated into the recipient language, giving rise to a new cognate class in the recipient if the recipient does not already have this trait. If the borrowed trait is already present in the recipient language, the borrowing event has no effect. First, we allowed borrowing to occur globally, with all

Figure 2. Demonstration of the SDM and the resulting binary output matrix. Data synthesis begins at the root of the input tree and moves towards the tips. The process is assumed to have been running infinitely in the past, and has generated the first language, containing trait 1. Trait 2 is born into this lineage. This is followed by a divergence event where the extant traits (1 and 2) are copied directly into both new lineages (Proto-AB and Proto-C). After this divergence event, trait 3 is born in the Proto-AB lineage, followed by the death of trait 1 in the top lineage (denoted by a cross). Trait 3 is then borrowed (denoted by a dotted arrow) into lineage C. The last divergence event leads to lineages A and B, where trait 4 arises in lineage B, and becomes borrowed into lineage A (dotted arrow). Finally, trait 2 dies in lineage A. At the end of the simulation, the lineages have the traits at the tips (e.g. lineage A has traits 3 and 4, while B has 2, 3 and 4). This can be recoded into a binary matrix as shown, to represent the presence or absence of the traits.

languages equally likely to receive a particular trait from any other contemporaneous language. This global borrowing scenario represents the somewhat unrealistic situation where all languages can borrow from each other. The second scenario we simulated was a ‘local borrowing’ constraint, where a language can borrow only from those lineages with which it shares a common ancestor within a specified time period. This makes the assumption that cultures are more prone to borrow from their neighbours and those that they have contact with (Nichols 1997).

We simulated data on both true topologies, under a global borrowing scenario and 1000- and 3000-year local borrowing scenarios. Within each of these scenarios, we simulated different rates of borrowing across a range expected to cover most feasible levels. Borrowing rates are relative to the death rate because the borrowing process increases rapidly at large rates. Values for the borrowing rate used in these simulations were 0 (i.e. no borrowing), 0.045, 0.224, 0.448, 0.672, 0.896, 1.344, 1.793 and 2.241. This translates to approximately 0, 1, 5, 10, 15, 20, 30, 40 and 50% of traits in a lineage being borrowed into another lineage per 1000 years. Note that, for a given language in the data, the proportion of observed traits that have entered its ancestry via borrowing is not simply related to the borrowing rate but is also dependent on the birth and death rates and the tree topology. To characterize the range of possible variation within each set of parameters, we replicated the analyses 10 times at a ‘realistic’ borrowing level of 15 per cent and at an ‘extreme’ level of 50 per cent. Analyses for other parameter settings were carried out once.

(c) Tree reconstruction

To assess the ability of phylogenetic methods to reconstruct the original input trees, we again used TRAITLAB (Nicholls & Gray 2006, 2008). TRAITLAB implements a Bayesian Markov chain Monte Carlo approach to reconstruct phylogenies and model parameters under the SDM without borrowing. Here, trees and model parameters are sampled from the space of all possible trees in proportion to their ability to explain the data under the SDM.

The death rate may be estimated as part of the fitting process. However, this requires that some constraints be imposed on the tree to provide time information. The constraints have the effect of narrowing the space of possible trees to only those trees that have the common divergence

point of the specified languages within a specified time period. Such a practice is common in Bayesian phylogenetic reconstruction of real linguistic data, where historical or archaeological data can inform us about the possible time frames in which the splitting of languages is likely to have occurred (Gray & Atkinson 2003; Gray *et al.* 2009). We arbitrarily selected two groupings in each of the true trees to act as constraints: one subtending 3–4 languages and the other subtending 9–11 languages. These nodes were constrained to within 10 per cent of their true ages (figure 1), approximating the variation in date estimates used elsewhere.

Each analysis was run for 5 million generations, which was long enough to ensure the chains had converged. Trees and parameters were sampled every 10 000th generation to reduce autocorrelation between samples. The first 3 million generations were discarded as ‘burn-in’ where the chain had yet to converge on the most likely region of tree space and the trees are unduly affected by the model priors. This resulted in 200 sampled trees per analysis.

To compare the estimated topologies with the input trees, we used the quartet distance metric (Day 1986; Steel & Penny 1993), as implemented in the program QUARTETDIST (Christiansen *et al.* 2006). This calculates the number of different combinations of four taxa (languages) in both trees and is normalized by dividing by the total number of quartets for the tree. The normalized quartet distance ranges from 0.0 when the two trees are identical to 1.0 when all quartets are different. This distance was calculated from the true tree to each of the post-burn-in sample trees from all simulated analyses. To assess the potential biasing effect of the constraints, we constructed two star (unresolved) topologies, constrained them in the same way and calculated the quartet distance between them and the true topologies. TRAITLAB was also used to estimate the age at the root of each post-burn-in sample tree.

3. RESULTS

The random trees have a normalized quartet distance of 0.703 (topology 1) and 0.662 (topology 2). If the normalized quartet distance of the estimated trees is between this value and 0 (a perfect match to the true tree), then the estimated topologies are closer to the true tree than predicted by chance. In what follows, we will take the

Figure 3. Mean quartet distances between the true tree topologies and the estimated topologies for both topology types across all borrowing levels ((a) tree 1 and (b) tree 2). The dotted line marks the quartet distance of the unresolved topology and the cross marks the mean quartet distance of the no-borrowing scenario ((i) local 1000, (ii) local 3000 and (iii) global).

percentage difference between the quartet distance of the estimated trees and the random trees as a measure of perturbation.

The normalized quartet distance (figure 3) shows that the trees recovered are extremely close to the true trees. Under the local borrowing scenarios, the analyses find trees very similar to the true trees, and the results are not markedly perturbed by the borrowing rates (average perturbation 0.71 and 7.58% for topologies 1 and 2, respectively). Topology 2 shows divergence from the input tree first, in the 3000-year local borrowing scenario (average perturbation = 14.58%), while topology 1 is more robust to these effects (average perturbation = 1.03%). The estimated topologies diverge more from the true topologies under the global borrowing scenarios, and this divergence increases relatively linearly with the amount of borrowing. However, even at their worst (perturbation in topology 1 = 53.22% and topology 2 = 56.61%), the quartet differences do not approximate of the random topologies (dotted lines in figure 3), demonstrating that the recovered trees are still recovering some of the true signal. The replicated analyses of the 15 and 50 per cent borrowing levels demonstrate mild variation between replicates (average variation in perturbation range in topology 1 = 10.37% and topology 2 = 13.74%). Overall, these results suggest that something very close to the true tree topology can be recovered by Bayesian phylogenetic methods, even under high levels of borrowing.

To investigate whether the inferences would be robust under a quite different model of cognate evolution, we replicated the 1000-year local borrowing scenario for topology 2 under a one-parameter model with gamma-distributed rate heterogeneity (Pagel & Meade 2004). The average perturbation across the entire range of borrowing levels was 2.15 per cent, suggesting that our results are robust to different models of reconstruction.

The differences between the age at the root of the true trees and those of the recovered trees (figure 4) show a general increase with borrowing level, where estimated root times become younger as borrowing increases. This

trend towards younger trees is unsurprising as borrowing increases the similarity between taxa, and therefore decreases branch lengths across the tree. This can be quantified by the percentage difference between the true age and the estimated root ages. The 1000-year local borrowing scenario is reasonably robust to this effect (percentage distance from true root age in topology 1 = 3.22% and topology 2 = 7.53%), with it becoming more noticeable in the 3000-year (5.22% topology 1, 32.54% topology 2) and global borrowing scenarios (41.80% topology 1, 39.23% topology 2). The global borrowing scenario shows a marked underestimation of the root time beginning at the 5 per cent borrowing level (1.31% topology 1, 6.43% topology 2). Topology 2 shows the same general patterns but with all estimates consistently underestimating the age of the tree. The greater difficulty in resolving topology 2's true age presumably reflects the difference in internal branch lengths between the trees. Topology 2 has shorter internal branches than topology 1. This makes the higher order structure in topology 2 harder to resolve, as well as increasing the probability that borrowing will collapse these nodes.

4. DISCUSSION

In this paper, we investigated the impact of horizontal transmission on Bayesian phylogenetic tree construction methods. Our results show that phylogenetic inference is remarkably robust to even quite high levels of borrowing. The crucial question is, of course, what scenarios of horizontal transmission are most realistic. Languages and cultures are able to borrow only from those they have contact with. For this reason, we would suggest that the global borrowing scenario simulated here is the least plausible. Instead, our local borrowing scenarios approximate the more likely situation where horizontal transmission occurs between geographical neighbours. To place the scenarios in more concrete terms, the 1000-year local borrowing scenario would be equivalent to allowing borrowing between most of the eastern Polynesian languages, such as Hawaiian and the Tuamotu Archipelago languages, or within the Slavic languages of

Figure 4. Mean reconstructed root time for each simulation under all borrowing scenarios ((a) tree 1 and (b) tree 2). The dotted line marks the true root age and the cross marks the root age under the no-borrowing scenario ((i) local 1000, (ii) local 3000 and (iii) global).

Indo-European. By contrast, the 3000-year local borrowing scenario would allow borrowing between all of the Oceanic languages, or between the Baltic and Slavic subgroups of Indo-European.

First, what levels of borrowing are plausible? Most phylogenetic analyses of linguistic data have used modified versions of the kind of basic vocabulary lists originally developed for lexicostatistics (Swadesh 1952). Items on the Swadesh lists are generally less susceptible to borrowing and evolve more slowly than less frequently used vocabulary. As an extreme example, more than 60 per cent of the total lexicon of English is known to be borrowed from French and Latin. By contrast, only 6 per cent of English basic vocabulary is borrowed from Romance languages and only 16 per cent borrowed in total (Embleton 1986). Most published language/culture phylogenies have followed a procedure of removing obvious borrowings before analysis. Therefore, we suggest that a plausible range of undetected borrowing may be between 0 and 20 per cent. We found that over the range of 0 to 30 per cent borrowing, there were very few differences between the true and estimated tree topologies. We have taken the value of 15 per cent borrowing to be representative of the higher end of this spectrum. At 15 per cent under the 1000-year local borrowing scenario, the average quartet difference between the true topologies and the estimated trees is extremely close to that of the true trees (average perturbation in topology 1 = 0.27% and topology 2 = 0.73%). In the 3000-year local borrowing scenario, the found trees are still very close to the true trees (topology 1 = 0.65%, topology 2 = 6.76%), with only the second topology showing a slight increase in overall difference. The quartet differences vary within the replicated analyses at the 15 per cent level and some replicates are less affected by borrowing. At the 15 per cent level, this perturbation ranges from 0.03–0.59% (topology 1) and 0.25–2.36% (topology 2) at 1000 years, to 0.18–1.44% (topology 1) and 2.29–20.42% (topology 2) at 3000 years. Overall, the estimation of tree topology is remarkably robust, especially in the realistic borrowing scenarios.

Second, in contrast to the topology results, the root time estimates are less robust to the effects of borrowing. These show a general shift towards younger ages as the amount of borrowing increases. However, at the 15 per cent level, the root time estimates are reasonably accurate under the 1000-year scenario (average difference to true root time in topology 1 = 2.18% and topology 2 = 6.13%). At the 3000-year local borrowing scenario, analyses of topology 1 again find the correct root ages (1.74% difference), while those of topology 2 show a shift to younger ages (28.31% difference). This has implications for the accurate phylogenetic dating of historical events. In general, ages will be underestimated as borrowing increases. This may or may not be a problem depending on the hypothesis being tested. For example, there has been considerable controversy over the use of phylogenetic methods to infer the age of the Indo-European language family (Gray & Atkinson 2003; Atkinson & Gray 2006; McMahon & McMahon 2006). Gray & Atkinson support an older farming-based dispersal from Anatolia *ca* 8500 years BP rather than the 'Kurgan' hypothesis that dates this family to *ca* 6000 years BP. Our results suggest that if unidentified borrowing has affected these divergence time estimates, then the real age may be older than that suggested by Gray & Atkinson, making the Kurgan hypothesis even less probable. It is worth noting that the more informative the constraints that can be added, the more robust the date estimates will be. In the simulations reported in this paper, we used only two date constraints, while Gray & Atkinson used fourteen.

Third, the shape of the true history is important. Our results show that the balanced tree (topology 1) is more robust to the effects of borrowing than the unbalanced tree (topology 2). This suggests that specific patterns of divergence can have important consequences on accurate reconstruction and inference. We suggest that topology 1 may be more resistant to the effects of borrowing because the internal branches are longer, while the shorter internal branches in topology 2 require a less conflicting signal to perturb them. Further research into the effects of

borrowing on different types of histories and speech community splitting events (e.g. Ross 1997) should elucidate this pattern further.

Fourth, the type of borrowing will affect the trees in different ways. This deserves some clarification, as some borrowing types may be more problematic than others. For the purposes of phylogenetic estimation, we can identify two major categories of borrowing: non-systematic and systematic. In non-systematic borrowing situations, cultures borrow a particular trait from other cultures they contact. This type of borrowing has also been called cultural borrowing (Bloomfield 1933). For example, English has borrowed the word 'taboo' from Tongan, but this borrowing has had little effect on the English language as a whole. As Goodenough (1997, p. 178) noted, the contact between America and Japan has led to borrowing of both culture and language in each direction, but has not diluted their 'distinctive phylogenetic identities'. In this situation, borrowing does not introduce any systematic bias, but increases the noise level in the data. If the majority of transmission is vertical, then the strongest signal in the dataset will still represent the true phylogeny. If the methods are sensitive enough, then the chance of inferring the true tree is high even under non-trivial levels of horizontal transmission. Such borrowing will still bias time–depth estimates towards younger ages by making the cultural inventories more similar.

By contrast, in systematic borrowing, traits flow predominantly from one culture to another, introducing systematic biases in the data. As an example, the Formosan language, Thao, has borrowed large quantities of lexicon from the neighbouring language Bunun (Blust 1996). The borrowed words primarily relate to women (including *binanau'az*, 'woman/wife') and other traditional female roles such as cooking and child rearing. This was probably a result of Thao men marrying Bunun women and acquiring their specialized vocabulary for this specific semantic domain (Blust 1996). In other situations, systematic borrowing can be even more prevalent and complex. The Oceanic language Yapese is a paradigmatic case of complex historical influence with at least five different sources of vocabulary: (i) early Palau, (ii) an unidentified non-Oceanic language, sources both (iii) inside and (iv) outside the Nuclear Micronesian language subgroup, and (v) relatively recent imports from Woleaian and Ulithian (Ross 1996). This systematic borrowing will perturb the topology by drawing the interacting languages together, making them appear to be more similar. This will have the further effect of making any time–depth estimates shallower. However, these types of borrowing tend to occur within a small subset of the taxa being examined and will not necessarily affect other parts of the tree or the broader scale inferences. Often such borrowing primarily affects a certain subset of the data, such as the 'women's work' subset in Thao. In such cases, it is more likely to be identified (Dutton 1995) and can therefore be removed.

This systematic/non-systematic distinction explains the conflict between our results and those of McMahon & McMahon (2005). McMahon and McMahon found that a level of 10 per cent borrowing caused disruption of the true topology in the majority of their analyses. However, a closer examination of their method reveals a number of key differences to our simulations. First, their simulation

allowed borrowing only between two of the terminal languages and only in the very final generations. This represents the type of borrowing we have called systematic. Second, the methods they used to recover the true tree are much less powerful—all the simulated data are converted directly to a distance measure before the tree search begins. This makes it impossible for the original sequence to be recovered and removes all the fine-grained information that likelihood-based Bayesian phylogenetic methods can use (Steel *et al.* 1988). Finally, McMahon and McMahon do not quantify either the uncertainty in their phylogenetic inferences nor whether this significantly affects inferences that could be made from the tree. In effect, the simulations by McMahon and McMahon estimate the effect of language hybridization on vintage lexicostatistics. It is unsurprising that this approach does not work well (Hoijer 1956; Blust 2000). By contrast, we have simulated the non-systematic type of borrowing, and used a Bayesian phylogenetic method to estimate trees.

To conclude, our results suggest that Bayesian phylogenetic methods are able to reliably estimate tree topology, even when confronted with substantial non-systematic borrowing. Root time estimates are less robust, showing a consistent trend towards younger ages as borrowing increases. However, this effect is not large at realistic borrowing levels. When there is very substantial horizontal transmission, there are strategies that can be used to untangle these complex histories (Gray *et al.* 2007). A method such as NeighborNet (Bryant & Moulton 2004; Bryant *et al.* 2005) that does not assume a tree-like history can be used to identify and visualize conflicting signal. These characters could then either be removed or explicitly taken into account with methods that can identify pattern heterogeneity, such as mixture models (Pagel & Meade 2004). Alternatively, the conflicting character histories could be decomposed into their own lineages and handled with a method such as reconciliation analysis (Page & Charleston 1998; Charleston 2003). In summary, while reticulation may be common in cultural evolution, it does not necessarily invalidate a phylogenetic approach.

We would like to thank David Bryant, Geoff Nicholls, Michael O'Brien, Robert Ross, David Welch and an anonymous reviewer for their discussion and suggestions.

REFERENCES

- Atkinson, Q. D. & Gray, R. D. 2005 Curious parallels and curious connections: phylogenetic thinking in biology and historical linguistics? *Syst. Biol.* **54**, 513–526. (doi:10.1080/10635150590950317)
- Atkinson, Q. D. & Gray, R. D. 2006 Are accurate dates an intractable problem for historical linguistics. In *Mapping our ancestors: phylogenetic approaches in anthropology and prehistory* (eds C. Lipo, M. J. O'Brien, M. Collard & S. Shennan), pp. 269–298. New Brunswick, NJ: Aldine Transactions.
- Atkinson, Q. D., Meade, A., Venditti, C., Greenhill, S. J. & Pagel, M. 2008 Languages evolve in punctuational bursts. *Science* **319**, 588. (doi:10.1126/science.320.5875.446a)
- Aunger, R. 2000 *Darwinizing culture: the status of memetics as a science*. Oxford, UK: Oxford University Press.
- Bateman, R., Goddard, I., O'Grady, R., Funk, V. A., Mooi, R., Kress, W. J. & Cannell, P. 1990 Speaking of forked

- tongues: the feasibility of reconciling human phylogeny and the history of language. *Curr. Anthropol.* **31**, 1–24. (doi:10.1086/203800)
- Bellwood, P. 1996 Phylogeny vs reticulation in prehistory. *Antiquity* **70**, 881–890.
- Bloomfield, L. 1933 *Language*. London, UK: Allen & Unwin.
- Blust, R. 1996 Some remarks on the linguistic position of Thao. *Oceanic Linguist.* **35**, 272–294. (doi:10.2307/3623175)
- Blust, R. 2000 Why lexicostatistics doesn't work: the 'universal constant' hypothesis and the Austronesian languages. In *Time depth in historical linguistics*, vol. 2 (eds C. Renfrew, A. McMahon & L. Trask), pp. 311–332. Cambridge, UK: McDonald Institute for Archaeological Research.
- Borgerhoff Mulder, M. 2001 Using phylogenetically based comparative methods in anthropology: more questions than answers. *Evol. Anthropol.* **10**, 99–111. (doi:10.1002/evan.1020)
- Bryant, D. & Moulton, V. 2004 Neighbor-Net: an agglomerative method for the construction of phylogenetic networks. *Mol. Biol. Evol.* **21**, 255–265. (doi:10.1093/molbev/msh018)
- Bryant, D., Filimon, F. & Gray, R. D. 2005 Untangling our past: languages, trees, splits and networks. In *The evolution of cultural diversity: phylogenetic approaches* (eds R. Mace, C. J. Holden & S. Shennan), pp. 67–84. London, UK: UCL Press.
- Carneiro, R. 2003 *Evolutionism in cultural anthropology: a critical history*. Boulder, CO: Westview Press.
- Charleston, M. A. 2003 Recent results in cophylogenetic mapping. *Adv. Parasitol.* **54**, 303–330. (doi:10.1016/S0065-308X(03)54007-6)
- Christiansen, C., Mailund, T., Pedersen, C. N. S., Randers, M. & Stissing, M. S. 2006 Fast calculation of the quartet distance between trees of arbitrary degrees. *Algorithms Mol. Biol.* **1**, 1–16. (doi:10.1186/1748-7188-1-16)
- Coward, F., Shennan, S., Colledge, S., Conolly, J. & Collard, M. 2008 The spread of Neolithic plant economies from the near east to northwest Europe: a phylogenetic analysis. *J. Archaeol. Sci.* **35**, 42–56. (doi:10.1016/j.jas.2007.02.022)
- Darwent, J. & O'Brien, M. J. 2006 Using cladistics to construct lineages of projectile points from northeastern Missouri. In *Mapping our ancestors: phylogenetic approaches in anthropology and prehistory* (eds C. Lipo, M. J. O'Brien, M. Collard & S. Shennan), pp. 185–208. New Brunswick, NJ: Aldine Transactions.
- Darwin, C. 1859 *On the origin of species*. London, UK: Murray.
- Day, W. H. E. 1986 Analysis of quartet dissimilarity measures between undirected phylogenetic trees. *Syst. Zool.* **35**, 325–333. (doi:10.2307/2413385)
- Dunn, M., Terrill, A., Reesink, G., Foley, R. A. & Levinson, S. C. 2005 Structural phylogenetics and the reconstruction of ancient language history. *Science* **309**, 2072–2075. (doi:10.1126/science.1114615)
- Dutton, T. 1995 Language contact and change in Melanesia. In *The Austronesians: historical and comparative perspectives* (eds P. Bellwood, J. J. Fox & D. T. Tryon), pp. 207–228. Canberra, Australia: Australian National University.
- Embleton, S. M. 1986 *Statistics in historical linguistics*. Bochum, Germany: Studienverlag Brockmeyer.
- Fortunato, L., Holden, C. J. & Mace, R. 2006 From bridewealth to dowry? A Bayesian estimation of ancestral states of marriage transfers in Indo-European groups. *Hum. Nat.* **17**, 355–376. (doi:10.1007/s12110-006-1000-4)
- Goodenough, W. H. 1997 Comment on social life in the Pacific. *Curr. Anthropol.* **38**, 177–178.
- Gould, S. J. 1987 *An urchin in the storm*. New York, NY: Norton.
- Gould, S. J. 1991 *Bully for brontosaurus*. New York, NY: Norton.
- Gray, R. D. & Atkinson, Q. D. 2003 Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* **426**, 435–439. (doi:10.1038/nature02029)
- Gray, R. D. & Jordan, F. M. 2000 Language trees support the express-train sequence of Austronesian expansion. *Nature* **405**, 1052–1055. (doi:10.1038/35016575)
- Gray, R. D., Greenhill, S. J. & Ross, R. M. 2007 The pleasures and perils of Darwinizing culture (with phylogenies). *Biol. Theory* **2**, 360–375. (doi:10.1162/biot.2007.2.4.360)
- Gray, R. D., Drummond, A. J. & Greenhill, S. J. 2009 Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* **323**, 479–483. (doi:10.1126/science.1166858)
- Greenhill, S. J. & Gray, R. D. 2005 Testing population dispersal hypotheses: Pacific settlement, phylogenetic trees and Austronesian languages. In *The evolution of cultural diversity: phylogenetic approaches* (eds R. Mace, C. J. Holden & S. Shennan), pp. 31–52. London, UK: UCL Press.
- Harmon, M. J., VanPool, T. L., Leonard, R. D., VanPool, C. S. & Salter, L. A. 2006 Reconstructing the flow of information across time and space: a phylogenetic analysis of ceramic traditions from prehispanic western and northern Mexico and the American southwest. In *Mapping our ancestors: phylogenetic approaches in anthropology and prehistory* (eds C. Lipo, M. J. O'Brien, M. Collard & S. Shennan), pp. 209–230. New Brunswick, NJ: Aldine Transactions.
- Harvey, P. H. & Pagel, M. 1991 *The comparative method in evolutionary biology*. Oxford, UK: Oxford University Press.
- Hoijer, H. 1956 Lexicostatistics: a critique. *Language* **32**, 49–60. (doi:10.2307/410652)
- Holden, C. J. 2002 Bantu language trees reflect the spread of farming across sub-Saharan Africa: a maximum-parsimony approach. *Proc. R. Soc. Lond. B* **269**, 793–799. (doi:10.1098/rspb.2002.1955)
- Holden, C. J. & Gray, R. D. 2006 Rapid radiation, borrowing and dialect continua in the Bantu languages. In *Phylogenetic methods and the prehistory of languages* (eds P. Forster & C. Renfrew), pp. 19–31. Cambridge, UK: McDonald Institute for Archaeological Research.
- Holden, C. J. & Mace, R. 2005 The cow is the enemy of matriliney: using phylogenetic methods to investigate cultural evolution in Africa. In *The evolution of cultural diversity: phylogenetic approaches* (eds R. Mace, C. J. Holden & S. Shennan), pp. 217–234. London, UK: UCL Press.
- Holden, C. J. & Shennan, S. 2005 How tree-like is cultural evolution? In *The evolution of cultural diversity: phylogenetic approaches* (eds R. Mace, C. J. Holden & S. Shennan), pp. 13–29. London, UK: UCL Press.
- Huelsenbeck, J. P. 1995 Performance of phylogenetic methods in simulation. *Syst. Biol.* **44**, 17–48. (doi:10.2307/2413481)
- Huelsenbeck, J. P. & Rannala, B. 1997 Phylogenetic methods come of age: testing hypotheses in an evolutionary context. *Science* **276**, 227–232. (doi:10.1126/science.276.5310.227)
- Jordan, P. & Shennan, S. J. 2003 Cultural transmission, language and basketry traditions amongst the California Indians. *J. Anthropol. Archaeol.* **22**, 42–74. (doi:10.1016/S0278-4165(03)00004-7)
- Jordan, F. M., Gray, R. D., Greenhill, S. J. & Mace, R. In press. Matrilocal residence is ancestral in Austronesian societies. *Proc. R. Soc. B*. (doi:10.1098/rspb.2009.0088)
- Laland, K. N. & Brown, G. R. 2002 *Sense and nonsense*. Oxford, UK: Oxford University Press.

- Mace, R. & Holden, C. J. 2005 A phylogenetic approach to cultural evolution. *Trends Ecol. Evol.* **20**, 116–121. (doi:10.1016/j.tree.2004.12.002)
- Mace, R. & Jordan, F. 2005 The evolution of human sex ratio at birth: a bio-cultural analysis. In *The evolution of cultural diversity: phylogenetic approaches* (eds R. Mace, C. J. Holden & S. Shennan), pp. 207–216. London, UK: UCL Press.
- Mace, R. & Pagel, M. 1994 The comparative method in anthropology. *Curr. Anthropol.* **35**, 549–564. (doi:10.1086/204317)
- McMahon, A. & McMahon, R. 2005 *Language classification by numbers*. Oxford, UK: Oxford University Press.
- McMahon, A. & McMahon, R. 2006 Why linguists don't do dates: evidence from Indo-European and Australian languages. In *Phylogenetic methods and the prehistory of languages* (eds P. Forster & C. Renfrew), pp. 153–160. Cambridge, UK: McDonald Institute for Archaeological Research.
- Mooers, A. O. & Heard, S. B. 1997 Inferring evolutionary processes from phylogenetic tree shape. *Q. Rev. Biol.* **72**, 31–54. (doi:10.1086/419657)
- Moore, J. H. 1994 Putting anthropology back together again: the ethnogenetic critique of cladistic theory. *Am. Anthropol.* **96**, 925–948. (doi:10.1525/aa.1994.96.4.02a00110)
- Moylan, J. W., Graham, C. M., Borgerhoff Mulder, M., Nunn, C. L. & Håkansson, N. T. 2006 Cultural traits and linguistic trees: phylogenetic signal in east Africa. In *Mapping our ancestors: phylogenetic approaches in anthropology and prehistory* (eds C. Lipo, M. J. O'Brien, M. Collard & S. Shennan), pp. 33–52. New Brunswick, NJ: Aldine Transactions.
- Neff, H. 2006 Archaeological-materials characterization as phylogenetic method: the case of copador pottery from Southeastern Mesoamerica. In *Mapping our ancestors: phylogenetic approaches in anthropology and prehistory* (eds C. Lipo, M. J. O'Brien, M. Collard & S. Shennan), pp. 231–248. New Brunswick, NJ: Aldine Transactions.
- Nicholls, G. K. & Gray, R. D. 2006 Quantifying uncertainty in a stochastic model of vocabulary evolution. In *Phylogenetic methods and the prehistory of languages* (eds P. Forster & C. Renfrew), pp. 161–171. Cambridge, UK: McDonald Institute for Archaeological Research.
- Nicholls, G. K. & Gray, R. D. 2008 Dated ancestral trees from binary trait data and its application to the diversification of languages. *J. R. Stat. Soc. B* **70**, 545–566. (doi:10.1111/j.1467-9868.2007.00648.x)
- Nichols, J. 1997 Modeling ancient population structures and movement in linguistics. *Annu. Rev. Anthropol.* **26**, 359–384. (doi:10.1146/annurev.anthro.26.1.359)
- O'Hara, R. J. 1988 Homage to Clio, or, toward an historical philosophy for evolutionary biology. *Syst. Zool.* **37**, 142–155. (doi:10.2307/2992272)
- Page, R. D. M. & Charleston, M. A. 1998 Trees within trees: phylogeny and historical associations. *Trends Ecol. Evol.* **13**, 356–359. (doi:10.1016/S0169-5347(98)01438-4)
- Pagel, M. 1999 Inferring the historical patterns of biological evolution. *Nature* **401**, 877–884. (doi:10.1038/44766)
- Pagel, M. & Meade, A. 2004 A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. *Syst. Biol.* **53**, 571–581. (doi:10.1080/10635150490468675)
- Pagel, M. & Meade, A. 2006 Estimating rates of lexical replacement on phylogenetic trees of languages. In *Phylogenetic methods and the prehistory of languages* (eds P. Forster & C. Renfrew), pp. 173–182. Cambridge, UK: McDonald Institute for Archaeological Research.
- Pagel, M., Atkinson, Q. D. & Meade, A. 2007 Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* **449**, 717–720. (doi:10.1038/nature06176)
- Penny, D., Hendy, M. D. & Steel, M. A. 1992 Progress with methods for constructing evolutionary trees. *Trends Ecol. Evol.* **7**, 73–79. (doi:10.1016/0169-5347(92)90244-6)
- Rexová, K., Frynta, D. & Zrzavý, J. 2003 Cladistic analysis of languages: Indo-European classification based on lexico-statistical data. *Cladistics* **19**, 120–127. (doi:10.1111/j.1096-0031.2003.tb00299.x)
- Rexová, K., Bastin, Y. & Frynta, D. 2006 Cladistic analysis of Bantu languages: a new tree based on combined lexical and grammatical data. *Naturwissenschaften* **93**, 189–194. (doi:10.1007/s00114-006-0088-z)
- Ross, M. 1996 Is Yapese Oceanic? In *Reconstruction, classification, description: festschrift in honor of Isidore Dyen* (ed. B. Nothofer), pp. 122–168. Hamburg, Germany: Abera.
- Ross, M. 1997 Social networks and kinds of speech-community events. In *Archaeology and language*, vol. 1 (eds R. Blench & M. Spriggs), pp. 209–261. London, UK: Routledge.
- Steel, M. & Penny, D. 1993 Distribution of tree comparison metrics—some new results. *Syst. Biol.* **42**, 126–141. (doi:10.2307/2992536)
- Steel, M., Hendy, M. D. & Penny, D. 1988 Loss of information in genetic distances. *Nature* **336**, 118. (doi:10.1038/336118a0)
- Swadesh, M. 1952 Lexico-statistic dating of prehistoric ethnic contacts. *Proc. Am. Philos. Soc.* **96**, 452–463.
- Tehrani, J. & Collard, M. 2002 Investigating cultural evolution through biological phylogenetic analyses of Turkmen textiles. *J. Anthropol. Archaeol.* **21**, 443–463. (doi:10.1016/S0278-4165(02)00002-8)
- Tëmkin, I. & Eldredge, N. 2007 Phylogenetics and material cultural evolution. *Curr. Anthropol.* **48**, 146–153. (doi:10.1086/510463)
- Terrell, J. E. 1988 History as a family tree, history as an entangled bank: constructing images and interpretations of prehistory in the South Pacific. *Antiquity* **62**, 642–657.
- Terrell, J. E., Kelly, K. M. & Rainbird, R. 2001 Foregone conclusions? In search of 'Papuan' and 'Austronesians'. *Curr. Anthropol.* **42**, 97–124. (doi:10.1086/318436)