



Supplementary Materials for

The Evolution of Antievolution Policies After *Kitzmiller v. Dover*

Nicholas J. Matzke

correspondence to: nick.matzke@anu.edu.au

This PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S11
Table S1

Other Supplementary Materials for this manuscript includes the following:

Supplemental Data S1 as a zipped archive of data and code.

Materials and Methods

Phylomemetics

Phylomemetics as a named field is young (6, 10), although applying phylogenies to cultural transmission, especially language (11, 12) has a long history (e.g., (13-15)). There are important outstanding questions about the degree to which cultural transmission can be modelled with phylogenies (13, 15, 16), and fundamental questions about the definition and coherence of the concept of “meme” (17, 18). Initial discussions of memes were dominated by population genetics analogies and by reductionism, selectionism and philosophical advocacy or criticism of these. Phylomemetics, on the other hand, takes a macroevolutionary perspective (meaning lineages and clades are the units of analysis) and heretofore has focused on data collection, phylogenetic estimation, and downstream inference. It thus may be an important step towards rigour in the study of memes, at least those that are large, mostly vertically transmitted, and slow-evolving enough to record phylogenetic history.

Because debates about memes and cultural phylogenies are unresolved, and because different cultural datasets may have different properties, in this study, special effort was put into tests for phylogenetic structure and correlations between character evolution and time. Such tests were originally devised for biological datasets, but are not often employed because phylogenetic signal is so obvious in most biological datasets (19-21).

Survey of Proposed Antievolution Policies and Their Strategies

Table S1 lists all of the policies (proposed or enacted) considered in this analysis. As a preliminary survey, the key strategies used in third-wave antievolutionism (2, 22, 23) are tabulated. The third-wave policies are compared to previous antievolution policies of historical importance in terms of major strategies used to denigrate or dilute the teaching of evolution. The major court decisions that rules particular antievolution policies unconstitutional are also listed in italics.

Explicitly religious strategies have been mostly abandoned in third-wave antievolutionism, as has the classic, but misguided, stratagem of forcing evolution to be taught as “theory, not fact.” However, a variety of strategies are still in circulation. Most of these can be directly traced to previous antievolution waves. For example, “academic freedom” was in fact explicitly invoked in the Balanced Treatment for Creation Science Acts of the 1980s, although they were nevertheless ruled unconstitutional. In addition, the Creation Science Legal Defense Fund, which defended the Balanced Treatment Acts, relabeled itself the Academic Freedom Legal Defense Fund soon after the 1987 *Edwards v. Aguillard* defeat (24).

The strategies of falsely suggesting that evolution is scientifically “controversial” and encouraging “critical analysis” are found in the disclaimer stickers that Alabama and Cobb County, Georgia required be stuck inside the front cover of any textbooks discussing evolution in the late 1990s and early 2000s, as well as in the Dover, PA

policy. Advocacy for teaching the “full range of scientific views” misappropriates a quote from the *Edwards v. Aguillard* decision in order to promote creationists’ incompetent criticisms of evolution, a strategy originating in the attempted “Santorum Amendment” to the 2001 No Child Left Behind Act. The amendment was drafted by ID advocate Phillip Johnson, but remained in the final Act only in non-binding report language (2). However, the tactic lives on in various third-wave policies.

Promotion of teaching the “strengths and weaknesses” of evolution, with the clear hope that teachers will teach creationist criticisms as “weaknesses,” traces back to language present in the Texas state science standards starting in the 1980s, until it was removed in contentious battles in 2008-2009. Antievolution policies are often adorned with disavowals of religious intent, and claims that the only intent is to protect those with alternative scientific views from discrimination, but such disavowals were common even in the creation science bills of the 1980s, and indeed, the “Balanced Treatment” strategy itself was merely an attempt to reframe the promotion of specific sectarian religious views in science classrooms as “fairness” to imagined scientific alternatives to evolution.

To summarize the texts used in this study, there are:

- 75 distinct antievolution policies in the Excel spreadsheet
- 73 policies, if two cases where there is a “teacher policy” and a distinct “faculty policy” in the same bill are collapsed to one policy each
- 71 legislative bills (including all available revisions); this excludes the Ouachita school district policy and the DI model policy
- 75 policy texts minus 8 excluded from the phylogeny for being too short for analysis equals 67 OTUs in the phylogeny (Operational Taxonomic Units, the equivalent of "species", i.e. tips in the tree)
- There are 60 distinct bill numbers. This excludes Ouachita, the DI Model Bill, and any legislative revisions.

Text Alignment

Bill texts were downloaded from the NCSE legislation news archive (7) or links therein. Text files (and sometimes PDFs) were saved for each bill, and the text files were edited to remove line breaks that did not represent paragraph breaks, line numbers, and non-ASCII characters. The text of each bill was pasted into the “bills_txt” worksheet of the “Academic_Freedom_Bills” Excel workbook, along with metadata, dates, and trait data about the strategies described in Table S1. These files are available in Supplemental Data.

The Discovery Institute Model Bill and the Ouachita Policy were broken into short segments – words or clauses – and used as templates against which to align the other bills. Bill alignment occurred in approximately the chronological order of bill publication. In some cases, a segment was broken up into two segments (adding a character) if it became clear that there were cases of the two segments being inherited

separately. Much more commonly, segments were added to the alignment when bills contained new features not found in the template policies. The names of these characters typically include the state and year when the segment was first observed.

Many bills had features that likely had non-phylogenetic causes – for instance, some legislatures mandate summaries, preface text such as “Be it enacted by the Legislature of the State of _____”, concluding implementation instructions, etc., and some do not. These were included in the text alignment for thoroughness and in order to help “frame” the alignment of meaningful characters, but these text segments were excluded during character coding. Similarly, some bills include large amounts of text on other issues; this text was not alignable with the main tradition and so was ignored in subsequent analysis.

Figure S1 shows a portion of the text alignment. This can be reproduced for any portion of the alignment by viewing the Excel file “align” worksheet, and zooming out.

Character Scoring

Once aligned, each text segment was treated as a character and scored 0/1 for absence/presence for each OTU (Operational Taxonomic Unit, a tip in the tree; equivalent to a species in a standard biological dataset). This created characters 1-111.

Further characters were scored by identifying variation within each “present” character state. The following procedure was used (conducted with an R script, but with major researcher involvement; this was not an attempt at “automated” character scoring).

1. All unique variants for that character were listed and counted
2. After manually examination of the unique variants, character strings unique to each variant that were likely to represent shared variation due to shared copying history were identified by hand (e.g., “biological evolution” and “biological origins” can be distinguished with the character strings “evolution” and “origins”).
3. The character strings were stored in the R script, along with the corresponding numerical characters state, and a grepl command was used to classify the character states for the character.
4. OTUs not matching any character strings were classified “?”, except in cases where presence/absence of a particular feature within a present segment was deemed likely to be informative. In these cases, a generic character string (typically “e” or “a”) was used to identify the “absent” or “other” state (typically state 0), and subsequent grepl searches overwrote with new character states when more specific character strings were found.
5. Autapomorphies were included, on the grounds that they might be parsimony-informative if new OTUs (or future bills) were added, and because they can be

mildly informative in likelihood/Bayesian analyses. (In addition, implementing the Mk-parsimony-informative ascertainment bias correction is much more difficult than implementing the Mk-variable correction. These corrections require calculating the likelihood for all unobservable character patterns. For Mk-variable, the number of unobservable patterns is just equal to the number of character states (the character has all 0s, all 1s, etc.). However, for Mk-parsimony-informative, the number of unobservable patterns increases dramatically with the number of OTUs and character states, so it is not clear that any available program actually can do the full calculation for highly multistate characters for many taxa.)

This procedure produced 307 additional characters, although they all had some missing data. These additional characters might be subject to the “absent tail/red tail/blue tail” criticism of Maddison (25). However, Maddison’s critique applies particularly to the situation where presence character states original multiple times, and coding additional variants within present characters can bias parsimony analyses. Here, most of the presence/absence characters in this analysis are very conserved (as documented below). In addition, Maddison notes that there is little alternative to the variants-when-present coding when there are multiple axes of variation in present characters; such a situation is often observed here. Finally, the focal analysis here is Bayesian rather than parsimony-based.

Volunteer Re-coding of Character Subset

There is some subjectivity involved in delineating and aligning segments, and in scoring characters. The input and output files are available in Supplemental Data so that future researchers may examine the judgment calls made. The effort made here was aimed at being thorough rather than absolutely exhaustive. Given that segment presence/absence is unambiguous, and that variants within present characters typically boiled down to just a few shared unique variants, it seems likely that an independent scoring attempt would capture essentially the same information, and produce very similar trees.

However, a reviewer suggested that a portion of the dataset be re-coded by an independent observer in order to check these assumptions. A subset of the text alignment (19 OTUs, characters 177-201, excluding some excluded from the final phylogenetic analysis) was extracted and pasted into an Excel worksheet. Four volunteer re-coders, recruited at the Society of Vertebrate Paleontology meeting (October 2015, Dallas) or the University of Tennessee, Knoxville, were asked to identify the character variants they saw for each text segment. The instructions they were given are available in the "recode_this" worksheet of the file "re-coding_texts.xlsx". Autapomorphies were allowed due to the small size of the subset matrix. Matzke’s coded characters using the full alignment were extracted and placed next to the re-codings in a summary worksheet ("summary_of_recoding").

The results of this procedure indicated that coding on the full dataset coded more character states, as expected given access to all of the OTUs; the recoders were given only a subset of the text alignment so as not to make the volunteer task onerous. However, most of the variations identified in the initial analysis were also identified by the re-coders. Some (rearrangements, text strings found in the full alignment but not the subset) could not be recognized by volunteers. Overall, comparison amongst coders indicates that the textual variants are fairly obvious upon inspection, and can be objectively identified. How exactly to code the textual variants can vary from coder to coder, but in the end, particularly over hundreds of characters, the same basic information (shared changes in texts, recording shared history) is likely to be detected.

Maximum Parsimony (MP) Analyses

After scoring, the R script converted the data into appropriate formats (NEXUS and TNT) for downstream analysis. Parsimony analyses were conducted in TNT version 1.1 (26) using custom R functions in TNTR (27) and BioGeoBEARS (28, 29). All TNT analyses were run using a version of the `aquickie.run` script included in the TNT download. The modifications to the script were cosmetic (removing pauses, and outputting statistics and tree files readable by TNTR). The analysis settings selected by the script (30) were to search for the optimal score 20 times independently, using function “`xmult`” on defaults, with 10 cycles of tree-drifting, settings described as “overkill” for a moderate-sized dataset. All runs produced a collection of most-parsimonious trees and a strict consensus tree; the latter was used for plotting and display of statistics (Bremer decay indices and bootstrap support frequencies; not run on null datasets).

Testing for Phylogenetic Signal

Phylogenetic signal will be completely destroyed by reshuffling each character among the OTUs (sampling without replacement). Inferring trees and statistics from datasets randomized in this way provides a null distribution that can be compared to observed values. Due to the slowness of TNT tree search when there is no tree structure, only 350 reshuffled datasets were used to generate the null. Observed values of tree length, Consistency Index (CI), Retention Index (RI), and Rescaled Consistency Index (RCI, equal to $CI \times RI$) were so far from the null expectation (fig. S2) that further generation of null datasets was judged unproductive. For example, the observed tree length of ~1000 steps is far below the tree lengths required to explain null datasets (2800-3800 steps). Technically, the use of 350 null datasets means P -values can go no lower than $1/350$ ($P < 0.0029$), but the distance between the observed and null values (many standard deviations) indicates that stating $P \ll 0.001$ is well-justified.

The same statistics were calculated for each individual character. Figure S3 shows the results for CI, but the plots for other statistics are similar (not shown). For comparison across characters, the plot uses normalized z -scores. The results confirm that most individual characters have dramatically higher congruence with the MP tree than expected by chance.

Testing if Fusion of Ouachita and AFA Texts Caused Detectable Character Incongruence

A fusion event, where text from a 2006 policy passed in Ouachita Parish, Louisiana became incorporated into the AFA tradition, has been noted by creationism observers (2), but it is worthwhile to ask if this event is likely to produce major anomalies in a phylogenetic analysis which assumes lineal inheritance.

This was assessed with an Incongruence Length Difference (ILD) test, where a character matrix is divided into two partitions, and the length of the MP tree inferred from the whole dataset is compared to the summed lengths of the two (possibly different) MP trees derived from the two partitions. The summed treelength is very likely to be smaller, so a null distribution is constructed by randomly constructing from the original dataset two partitions of the same size as the original partitions. If the observed summed treelength is shorter than the null, this is taken as evidence against the null hypothesis that the two datasets derive from the same topology.

ILD has received criticism as a non-conservative test of either reticulation or dataset incongruence, as it is prone to false positives: topological incongruity can be wrongly inferred when, instead, the partitions are evolving under different rates or models (31). A significant rejection of the null therefore may indicate only a rejection of perfect homogeneity between the partitions. Few datasets outside of simulations are likely to be perfectly homogenous, so ILD results should be interpreted with caution.

Nevertheless, comparing the tree lengths for observed and randomized partitions gives some measure of whether or not partitioning dramatically shortens trees. Partitions were constructed using characters classified as “AFA-typical” (chatype a in the “align” worksheet) or “SEA-typical” (chatype b or ab). A null distribution was constructed by randomly assigning characters to two partitions of the same size 1000 times and inferring MP trees from these datasets. The results are shown in Figure S4. The observed tree length under two partitions (985 steps) is in the lower part of the null distribution ($P \sim 0.102$), suggestive of moderate inhomogeneity but not a radical difference in the tree topology or model between the partitions.

Inspection of MP trees from AFA and SEA subsets of the OTUs (figs. S5-S6) with the MP tree from the full dataset (fig. S7) indicates that while the subset datasets infer trees with good correlation between the character change and time, the full dataset tends to infer a similar topology for the SEA bills, but places the root of the SEA subtree (where the AFA phylogeny connects to the SEA phylogeny) away from Ouachita. This is probably due to Ouachita scoring absent for many AFA characters, leading the analysis to attach the AFAs to branches sharing more AFA characters.

As Ouachita is just one OTU, rather than an entire clade with a different evolutionary history (which might be treated with a network analysis), it was judged preferable to include it as an OTU in the final phylogenetic analyses, in order to correctly date the origin of Ouachita-derived character states. Expert knowledge (2) and the dating

support placing a prior constraint on the SEA subtree of (Ouachita,(other SEAs)), with the AFAs being outgroups with respect to the AFAs. This constraint was enforced in the Bayesian tip-dating analysis.

Testing the Reasonableness of a “Morphological Clock” Model

In paleontological tip-dating analyses, the “morphological clock” is a model which states that changes in discrete morphological character states occur at some constant average rate. “Relaxed clock” models allow the rate to vary on different branches, although the assumption of a constant overall average rate is maintained. Of course, any collection of characters evolving on a tree must, mathematically, have an average rate. The key question with clock models is whether or not having one average rate throughout the study time period is a poor model that misleads dating, as might be the case with (a) a rapid radiation after a mass extinction, or (b) wild shifts in rate throughout the tree, such that time and morphological change do not correlate.

Here, the data are discrete characters coded from texts rather than morphology, but the “morphological clock” terminology is retained for consistency with the tip-dating literature. The reasonableness of the clock assumption was assessed by linear regression of time-above-root for the MP AFA and SEA subtrees, against the number of parsimony steps inferred on the MP parsimony tree from root-to-tip for each subtree. These data were also plotted. Statistically significant positive slope indicates significant positive correlation between time elapsed and amount of morphological change.

Tip-dating with Sampled Ancestors

A Bayesian tip-dating analysis was set up in Beast 2.1.3 (32). “Tip-dating” uses non-contemporaneous tips and a morphological clock to estimate dates, in contrast to the traditional “node-dating” technique, which relies on informative prior calibrations of the dates of certain nodes (33-35). Rather than BEAUTi, the analysis was set up using the R package BEASTmaster (36), which streamlines the process of converting a NEXUS character matrix, tip dates, and priors into a Beast2 XML setup for a tip-dating analysis. The R scripts, Excel settings input file, and XML output are available in Supplemental Data.

A relaxed morphological clock, with lognormally distributed branch rate variation and flat priors, was used. The site model was Markov- k , all characters unordered and sharing a common clock rate, but with gamma-distributed variation among sites, with 4 rate categories. The Mk-variable ascertainment bias correction was used, for reasons described above. The tree model was a birth-death process with serial sampling (BDSS), with the possibility of sampling direct ancestors (8, 9) (BDSS-SA, called SABD in the Beast2 and BEASTmaster).

Allowing the DI model bill to have a wide date distribution made convergence more difficult, so four independent runs were conducted, sampling every 5000 generations for more than 50 million generations per run. After confirming that each run reached

stationarity and convergence in log-probability and parameter estimates, the post-burnin runs were merged using burntrees (37). All parameters had ESS (estimated sample size) >100, typically >1000.

Application of the birth-death-serial sampling model to the phylomemetics of legislation requires careful consideration. While speciation has a ready interpretation (bill copying), and sampling represents publication, the interpretation of extinction in the case of an evolving legislative tradition is more difficult. The death of a bill in a legislature is not an appropriate interpretation, since even a legally “dead” bill can still serve as a text source for a subsequent bill. A possible interpretation of extinction can be derived from the observation that it appears that legislators mostly copy their legislation from proposals of the previous few years. In that sense, a bill text lineage is “alive” as long as it is still being kept in mind as a source by legislators, and it is “extinct” once they have forgotten about it and moved on to newer text sources.

Inference of direct ancestors was conducted by enumerating the frequency of terminal branches with 0 time length in the post-burnin MCMC collection of trees. In essence, these are OTUs with a date that fits the date expected under the relaxed morphological clock, and that have no character states that give high probability for placing the OTU in a more derived position.

The phylogeny in Figure 1 is the Maximum Clade Credibility tree calculated from the posterior distribution of trees using TreeAnnotator. Figure S9 diagrams uncertainty in dating estimates (95% HPD, highest posterior density of dates).

Character Mapping

The branch colors in Figure 1 are from a parsimony ancestral state reconstruction done in Mesquite, using “policy targeted” as a multistate unordered trait. Similar ancestral state reconstructions have been done for all characters and traits, and may be viewed by opening the NEXUS files in Supplemental Data in Mesquite (files “traits_v1_mapped_on_tree_67_taxa.nex” and “morph_v1_mapped_on_tree_67_taxa.nex”. (Here, by “character” I mean characters used to infer the phylogeny, and “traits” to mean other features of interest that were not used to infer the phylogeny, such as state of origin. There is some overlap between the two, as some traits were assembled by combining characters.)

Tree Balance and Diversification-shift Analysis

The “imbalance” or “asymmetry” of a phylogeny can be an indicator of the processes that produced it. “Balanced” or “symmetrical” trees have relationships like ((A,B),(C,D)), while “imbalanced” or “asymmetrical” trees have relationships like (A,(B,(C,D))).

Visual inspection of Figure 1 suggests strong asymmetry. This is confirmed by calculating Colless’ index on the MCC tree (Colless’ index=9.52, under an equal-rates

Markov (ERM) null model, $P \sim 0$) with the R package *apTreeShape* (38). This means there was a strong bias in which lineages were selected for further copying, supporting the suggestion that “success sells” in the process of copying and promoting third-wave antievolution legislation.

Inference of shifts in diversification rate are now common in macroevolutionary analyses, and it is of interest to see if such shifts can be detected in a phylomemetic case. However, available phylogenetic comparative methods for diversification are relatively limited for “paleontological” time trees: most models condition on all the tip taxa being extant (in other words, most methods assume an ultrametric tree, i.e. a dated tree with all of the tips living in the present). However, diversification on paleontological timetrees can be studied using the Δ_1 shift statistic calculated as the tree is pruned by time-slice.

The Δ_1 shift statistic (39) is a topology-based statistic that compares the likelihood ratio for the difference in diversity of two sister clades in an ingroup, with the same ratio calculated for the ingroup versus the outgroup, all in comparison to a null equal-rates Markov (ERM) expectation. This procedure aims to distinguish diversification shifts at the node of interest from diversification shifts elsewhere in the tree. Detailed consideration of the use of topology-based diversification-shift statistics in paleontological timetrees (40-42) recommends pruning the tree by time-slice (or, in reverse, “growing” the tree) and calculating the shift statistic at all nodes for all time steps. Nodes that consistently show low p-values are most likely to have experienced shifts in diversification. This procedure was implemented using the *apTreeshape* function *shift.test* to calculate Δ_1 (38), and *APE* (43) and *BioGeoBEARS* time-stratification functions (28, 29) such as *chainsaw* to prune the tree and plot the shift *P*-values. Time slices were performed every 2 years (because many legislatures operate on a biannual cycle). The same nodes were highlighted except in the most-pruned trees (height 2 and 4), when the tree was too small to yield substantial evidence against the null model.

This analysis identified three nodes with diversification shifts unlikely to be explained ($P < 0.1$) by the null hypothesis of no change in rates. Although in all cases $0.05 < P < 0.1$ and so statistical significance is debatable, support for diversification shifts was judged “substantial” (40) on the grounds that topology-based tests are conservative, dataset size is unavoidably limited, and the indicated shifts each occur at nodes where the right branch has a single OTU and the left branch contains the entire rest of the tradition. The nodes that consistently showed diversification rate shifts are indicated with stars in Fig. S10. Notably, inferred diversification shifts tend to be closely associated with direct ancestors independently inferred in the BDSS-SA analysis, confirming that each biennium, one or a few selected bill texts tended to be chosen for copying in the new legislative session.

Supplementary Text

Supplemental Results

An overview of the results of the diversification shift analysis, inferring the date of the DI model bill, the parsimony analysis and the time/morphological change correlation analysis is shown in Fig. S10. A plot of the inferred rate of morphological change on the tree (another product of the BEASTmaster/Beast2 tip-dating analysis) is also shown.

Creationist Motivations of Third-Wave Antievolution Legislation

The National Center for Science Education online archive of AFAs/SEAs contains a summary of contemporary news reports for each bill. These often include quotes from sponsors of the bills. Quotes from these summaries, and links to the NCSE summaries, are provided in the worksheet “Religious_purpose” of the file “Academic_Freedom_bills.xlsx”. Previous court decisions have ruled that creationism and intelligent design constitute particular religious views, and court decisions on the creationism/evolution issue have relied heavily on the *Lemon* test (derived from *Lemon v. Kurtzman*, 1971). The most important prong of the *Lemon* test states that if a government policy has the purpose of government endorsement of a particular religious view, it constitutes an unconstitutional violation of the First Amendment's prohibition against government establishment of religion. Therefore, if sponsors, authors, or testimony in favor of a “third wave” bill have favored creationism or intelligent design in legislation, or are activists for such organizations, this is scored as evidence of religious purpose, as are antievolution phrases common with creationists (e.g., “theory not fact”). More general statements indicating concern that evolution undermines religion are also counted. Not counted is mere affiliation with a religious group, or generic affiliation with conservative groups. In many cases, the same sponsor or co-sponsor is on many bills over several years; in such situations, quotes may be repeated.

For at least 63 of the 65 bill texts used the phylogeny (note that this count includes multiple versions of the same bill, but excludes the Ouachita and DI model policies), documentation of sponsor statements indicating creationist motivation were found.

Supplemental Bill Text History

After the analysis was complete, it came to my attention that some of the text of the Ouachita Policy is known to have been copied from a Proposed School Board Policy promulgated by retired military judge Darrell White (44, 45). White is a member of Louisiana Family Forum (a group historically involved in many creationist efforts in that state) and a “lifetime member of the Creation Museum” (45). He was also a public advocate for the passage of the Ouachita policy and Louisiana’s SEA. White’s policy can be seen in a February 8, 2005 archive of the (now defunct) website judgewhite.com (<https://web.archive.org/web/20050208075130/http://judgewhite.com/docs/proposedresolution.pdf>). An even earlier copy is appended to an “Open Letter to LA Educators” (dated March 9, 2004) by Darrell Scott, a parent of a victim of the Columbine High School shootings. The letter expresses concern that the teaching of evolution leads to atheism and school shootings (<https://web.archive.org/web/20050210161118/http://judgewhite.com/docs/dscottletter.pdf>).

The Judge White text is clearly the source for some of the Ouachita policy, and includes common Discovery Institute talking points, although it does not have some key pieces like the targeting of human cloning and global warming in addition to evolution.

The Discovery Institute website academicfreedompetition.com has recently posted a “2015-2016” version of an Academic Freedom bill (<http://www.academicfreedompetition.com/freedom.php>, “Model Academic Freedom Bill”). Despite the “Academic Freedom” label, this model bill is clearly in the SEA tradition, complete with the phrase “teaching of some scientific subjects, such as biological evolution, the chemical origins of life, global warming, and human cloning, can cause controversy.”

The original DI Model Bill seems to have disappeared from the website, but a 2012 copy of the version posted in 2007 can still be seen at <https://archive.is/20120717194631/http://www.academicfreedompetition.com/freedom.php>.

Supplemental Bill Text History: 2010 NCSE in-house analysis

Part of the inspiration for this study was a graphic generated at NCSE in 2010, depicting the evolution of the legislative tradition up to that point (Figure S11). This graphic was not published, but it was on occasion used in lectures and shown to visitors. I interpret the graphic as equivalent to phylogenies drawn by hand in pre-cladistic analyses, using a combination of expert knowledge, knowledge of the fossil record, and intuition. The analysis has many similarities to the conclusions drawn here (for example, the fusion of the Ouachita policy into the AFA tradition), but also some differences (for example, the present analysis does not identify the DI Model Bill as directly ancestral to the main tradition).

Supplemental Acknowledgements

I would like to thank the National Center for Science Education for its online legislation news archive, and acknowledge helpful comments or discussion from Eugenie Scott, Glenn Branch, Josh Rosenau, Chris Knight, Eric Meikle, and Joe Felsenstein; and to Sandy Kawano, Julien Louys, Michelle Lawing, and Nathan Jackson for volunteering to independently re-code a character subset. Portions of this work (including BEASTmaster and BioGeoBEARS development) were supported by NSF Award 0919124, and by the National Institute for Mathematical and Biological Synthesis (NIMBioS), an Institute sponsored by the National Science Foundation through NSF Award #DBI-1300426 with additional support from The University of Tennessee, Knoxville. Work was also funded by the Australian Research Council’s Discovery Early Career Researcher Award #DE150101773, and by The Australian National University.

Supplemental References

1. C. A. Bleckmann, Evolution and Creationism in Science: 1880–2000. *Bioscience* **56**, 151-158 (2006).
2. G. Branch, E. C. Scott, J. Rosenau, Dispatches from the Evolution Wars: Shifting Tactics and Expanding Battlefields. *Annual Review of Genomics and Human Genetics* **11**, 317-338 (2010).
3. N. J. Matzke, in *But Is It Science? The Philosophical Question in the Creation/Evolution Controversy, Updated Edition*, M. Ruse, R. T. Pennock, Eds. (Prometheus Books, Amherst, New York, 2009), chap. 20, pp. 377-413.
4. K. Padian, N. J. Matzke, Darwin, Dover, 'Intelligent Design' and textbooks. *The Biochemist* **31**, 8-22 (2009).
5. Mississippi HB214, signed into law in 2006, includes an "academic-freedom"-style amendment added to an unrelated bill, but the language is too short (just one sentence) to include in the phylogeny.
6. C. J. Howe, H. F. Windram, Phylomemetics—Evolutionary Analysis beyond the Gene. *PLoS Biol.* **9**, e1001069 (2011).
7. National Center for Science Education, "Academic Freedom" Bills by State & Year. <http://ncse.com/creationism/general/academic-freedom-bills-by-state-year> (2015).
8. A. Gavryushkina, D. Welch, T. Stadler, A. J. Drummond, Bayesian Inference of Sampled Ancestor Trees for Epidemiology and Fossil Calibration. *PLoS Comput Biol* **10**, e1003919 (2014).
9. A. Gavryushkina *et al.*, Bayesian total evidence dating reveals the recent crown radiation of penguins. *arXiv*, (2015).
10. J. J. Tehrani, The Phylogeny of Little Red Riding Hood. *PLoS ONE* **8**, e78871 (2013).
11. R. D. Gray, A. J. Drummond, S. J. Greenhill, Language Phylogenies Reveal Expansion Pulses and Pauses in Pacific Settlement. *Science* **323**, 479-483 (2009).
12. R. Bouckaert *et al.*, Mapping the Origins and Expansion of the Indo-European Language Family. *Science* **337**, 957-960 (2012).
13. N. Eldredge, in *Perspectives in Ethology*, F. Tonneau, N. Thompson, Eds. (Springer US, 2000), vol. 13, chap. 5, pp. 113-153.
14. N. Eldredge, An overview of piston-valved cornet history. *Historic Brass Society Journal* **14**, 337-390 (2002).
15. M. J. O'Brien, M. Collard, B. Buchanan, M. T. Boulanger, Trees, thickets, or something in between? Recent theoretical and empirical work in cultural phylogeny. *Israel Journal of Ecology & Evolution* **59**, 45-61 (2013).
16. N. Eldredge, in *Macroevolution in Human Prehistory*, A. Prentiss, I. Kuijt, J. C. Chatters, Eds. (Springer New York, 2009), chap. 12, pp. 297-316.
17. M. Midgley, in *Alas, poor Darwin: Arguments against evolutionary psychology*, H. Rose, S. Rose, Eds. (Harmony Books, New York, 2000), pp. 67-84.
18. C. Coker, War, memes and memplexes. *Int. Affairs* **84**, 903-914 (2008).
19. D. Penny, L. R. Foulds, M. D. Hendy, Testing the theory of evolution by comparing phylogenetic trees constructed from five different protein sequences. *Nature* **297**, 197-200 (1982).

20. D. Penny, M. D. Hendy, M. A. Steel, in *Phylogenetic Analysis of DNA Sequences*, M. M. Miyamoto, J. Cracraft, Eds. (Oxford University Press, New York, 1991), pp. 155-183.
21. D. Penny, M. D. Hendy, Testing methods of evolutionary tree construction. *Cladistics* **1**, 266-278 (1985).
22. E. C. Scott, *Evolution versus Creationism: An Introduction*. (2009).
23. N. Matzke, The Evolution of Creationist Movements. *Evo Edu Outreach* **3**, 145-162 (2010).
24. R. Moore, *Evolution in the courtroom: A reference guide*. (ABC-CLIO, 2002), pp. 383.
25. W. P. Maddison, Missing data versus missing characters in phylogenetic analysis. *Syst. Biol.* **42**, 576-581 (1993).
26. P. A. Goloboff, J. S. Farris, K. C. Nixon, TNT, a free program for phylogenetic analysis. *Cladistics* **24**, 774-786 (2008).
27. N. J. Matzke, TNTR: R functions to aid analyses in the cladistics program TNT. <http://phylo.wikidot.com/tnt> (2015).
28. N. J. Matzke, Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Syst. Biol.* **63**, 951-970 (2014).
29. N. J. Matzke, BioGeoBEARS: an R package for model testing and ancestral state estimation in historical biogeography. <http://phylo.wikidot.com/biogeobears> (2015).
30. P. A. Goloboff, Analyzing large data sets in reasonable times: solutions for composite optima. *Cladistics* **15**, 415-428 (1999).
31. F. K. Barker, F. M. Lutzoni, The utility of the incongruence length difference test. *Syst. Biol.*, 625-637 (2002).
32. R. Bouckaert *et al.*, BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Comput Biol* **10**, e1003537 (2014).
33. R. A. Pyron, Divergence time estimation using fossils as terminal taxa and the origins of Lissamphibia. *Syst. Biol.*, syr047 (2011).
34. M. A. Alexandrou, B. A. Swartz, N. J. Matzke, T. H. Oakley, Genome duplication and multiple evolutionary origins of complex migratory behavior in Salmonidae. *Mol. Phylogen. Evol.* **69**, 514-523 (2013).
35. H. M. Wood, N. J. Matzke, R. G. Gillespie, C. E. Griswold, Treating Fossils as Terminal Taxa in Divergence Time Estimation Reveals Ancient Vicariance Patterns in the Palpimanoid Spiders. *Syst. Biol.* **62**, 264-284 (2013).
36. N. J. Matzke, BEASTmaster: Automated conversion of NEXUS data to BEAST2 XML format, for fossil tip-dating and other uses. <http://phylo.wikidot.com/beastmaster> (2015).
37. J. Nylander, burntrees. <https://www.abc.se/~nylander/burntrees/burntrees.html> (2015).
38. N. Bortolussi, E. Durand, M. Blum, O. Francois, apTreeshape: Analyses of Phylogenetic Treeshape. <http://CRAN.R-project.org/package=apTreeshape> (2012).
39. B. R. Moore, K. M. Chan, M. J. Donoghue, in *Phylogenetic Supertrees*. (Springer, 2004), pp. 487-533.

40. G. T. Lloyd *et al.*, Dinosaurs and the Cretaceous terrestrial revolution. *Proceedings of the Royal Society of London B: Biological Sciences* **275**, 2483-2490 (2008).
41. M. Ruta, D. Pisani, G. T. Lloyd, M. J. Benton, A supertree of Temnospondyli: cladogenetic patterns in the most species-rich group of early tetrapods. *Proceedings of the Royal Society of London B: Biological Sciences* **274**, 3087-3095 (2007).
42. J. E. Tarver, P. C. Donoghue, The trouble with topology: phylogenies without fossils provide a revisionist perspective of evolutionary history in topological analyses of diversity. *Syst. Biol.* **60**, 700-712 (2011).
43. E. Paradis, J. Claude, K. Strimmer, APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289-290 (2004).
44. B. Forrest, The Discovery Institute, the LA Family Forum, and the "LA Science Education Act" UPDATED. *Talk to Action*, (2008).
<http://www.talk2action.org/story/2008/6/26/18920/8497>
45. Z. Kopplin, Dismissing Darwin. *Slate*, (2015).
http://www.slate.com/articles/health_and_science/science/2015/04/creationism_in_louisiana_public_school_science_classes_school_boards_and.html

Fig. S1. Zoomed-out view of a portion of the manual text alignment. The full alignment may be viewed in the Excel file in Supplemental Data. Red: text originating in the AFA tradition. Green: text originating in the Ouachita/SEA tradition. The first and fourth columns represent the Ouachita and DI policies, respectively. The two highlighted columns are the SEA bills passed in Louisiana and Tennessee, respectively.

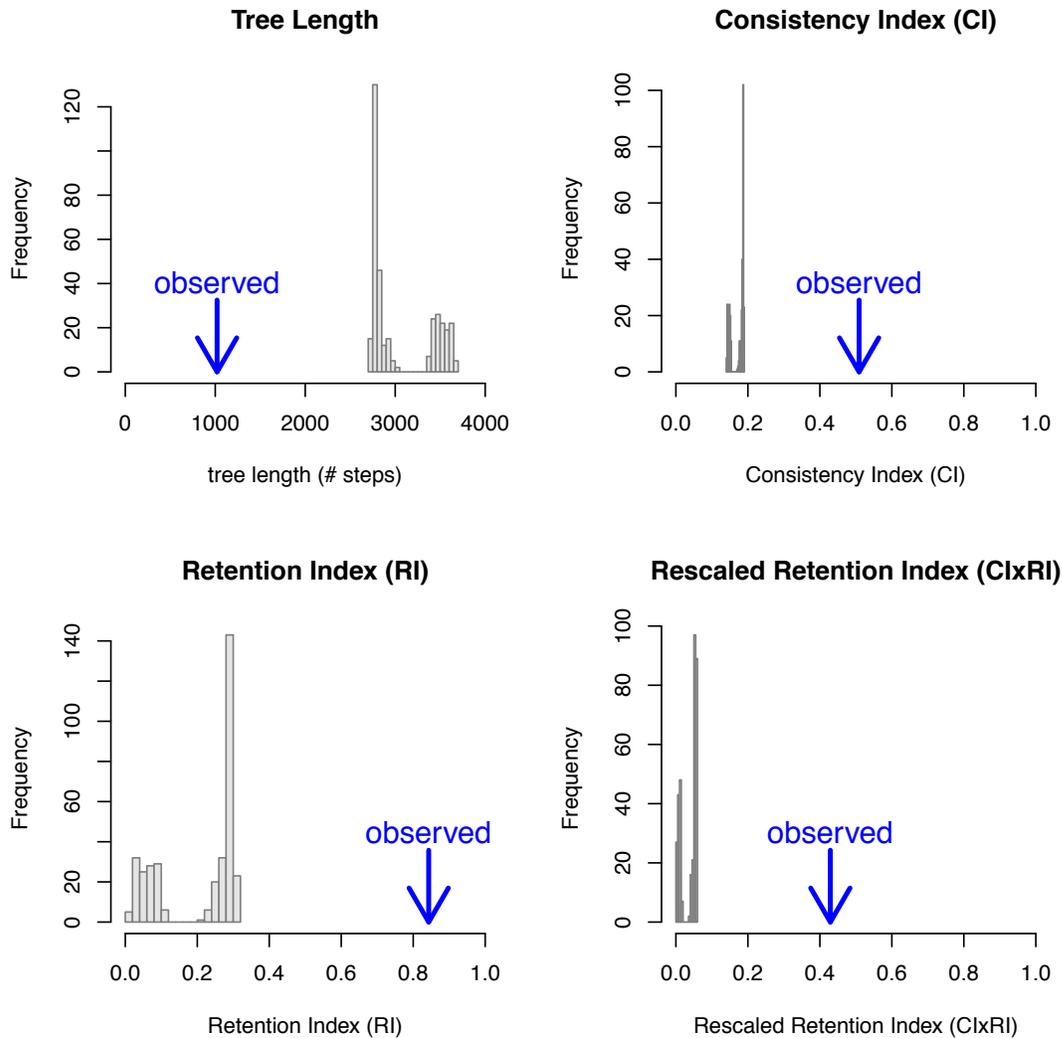


Fig. S2. Observed values of standard statistics for the maximum parsimony (MP) cladograms estimated from the legislation character matrix, using TNT and a lightly modified version of the `aquickie.run` script. Tree topology was unconstrained. The null distributions for the statistics (grey histograms) were constructed by reshuffling each character among the tips and estimating a cladogram from the new dataset. This was repeated 350 times. The bimodality in the null distribution probably just indicates the difficulty of finding MP trees using a character matrix with no signal -- the MP search had to be limited for the null datasets due to the tendency of the `aquickie` settings to find thousands of trees of equal length and crash the program. Regardless of which peak in the null is preferred, the empirical p-value of the observed statistics is certainly less than $p=1/350=0.00286$, and likely far below that ($p \ll 0.001$).

Observed and null z-scores for CI

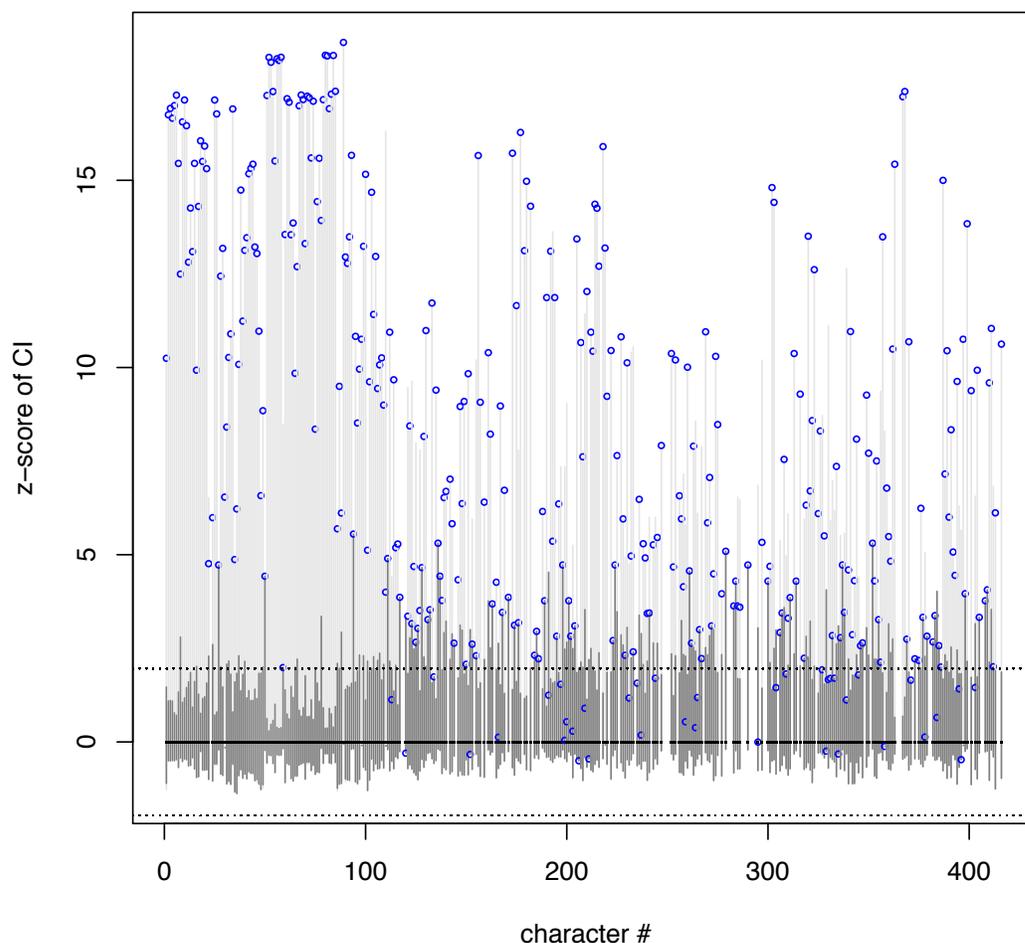


Fig. S3. Observed values (blue dots) and the 95% confidence intervals (dark grey bars) of the z-score of Consistency Index (CI) for individual characters. The cladograms and null distributions are the same as described in Fig. S2. The z-score represents the number of standard deviates above or below the mean. The dotted black lines indicates the theoretical 95% confidence interval (± 1.96) of the standard normal distribution, but this is not always met exactly by the null distribution for each character, due to the non-normality of the null distribution and individual variation amongst characters, percentage of missing data, etc., so this plot should be regarded as heuristic. Nevertheless, most characters exhibit far greater consistency than would be expected by chance, with the presence/absence characters (#1-111) showing particularly high consistency ($z > 10$ for most characters). Autapomorphic characters produce blanks in the plot, but were retained in the analysis to preserve consistency of numbering, and because they can be informative in the Bayesian analysis.

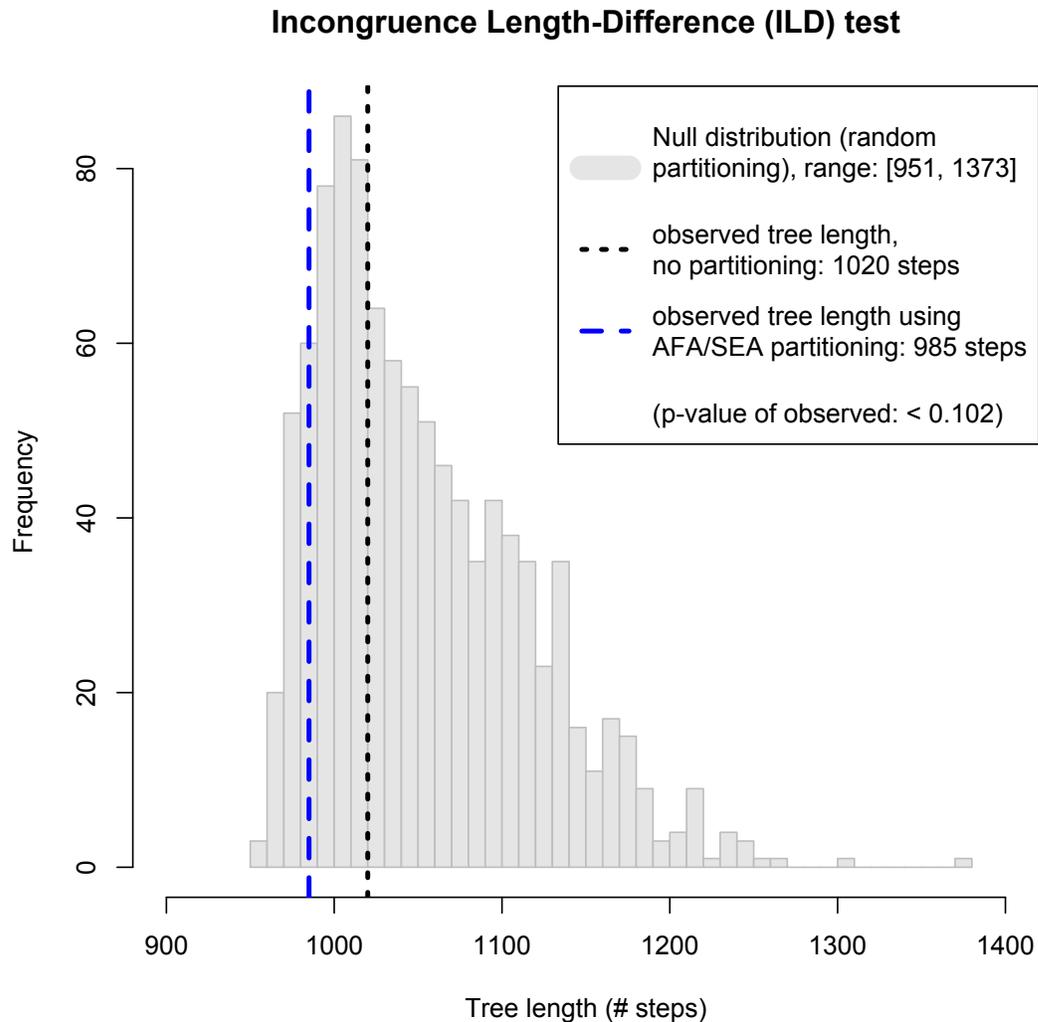


Fig. S4. Observed value and null distribution for the Incongruence Length Difference (ILD) test. For the observed value, the characters were partitioned into AFA- and SEA-derived characters, trees were inferred from each partition, and their lengths were summed. For the null distribution, the same procedure was followed, with partitions of the same size selected, but with the characters randomly permuted between the partitions 100 times.

**MP tree
(bootstrap support/frequency)**

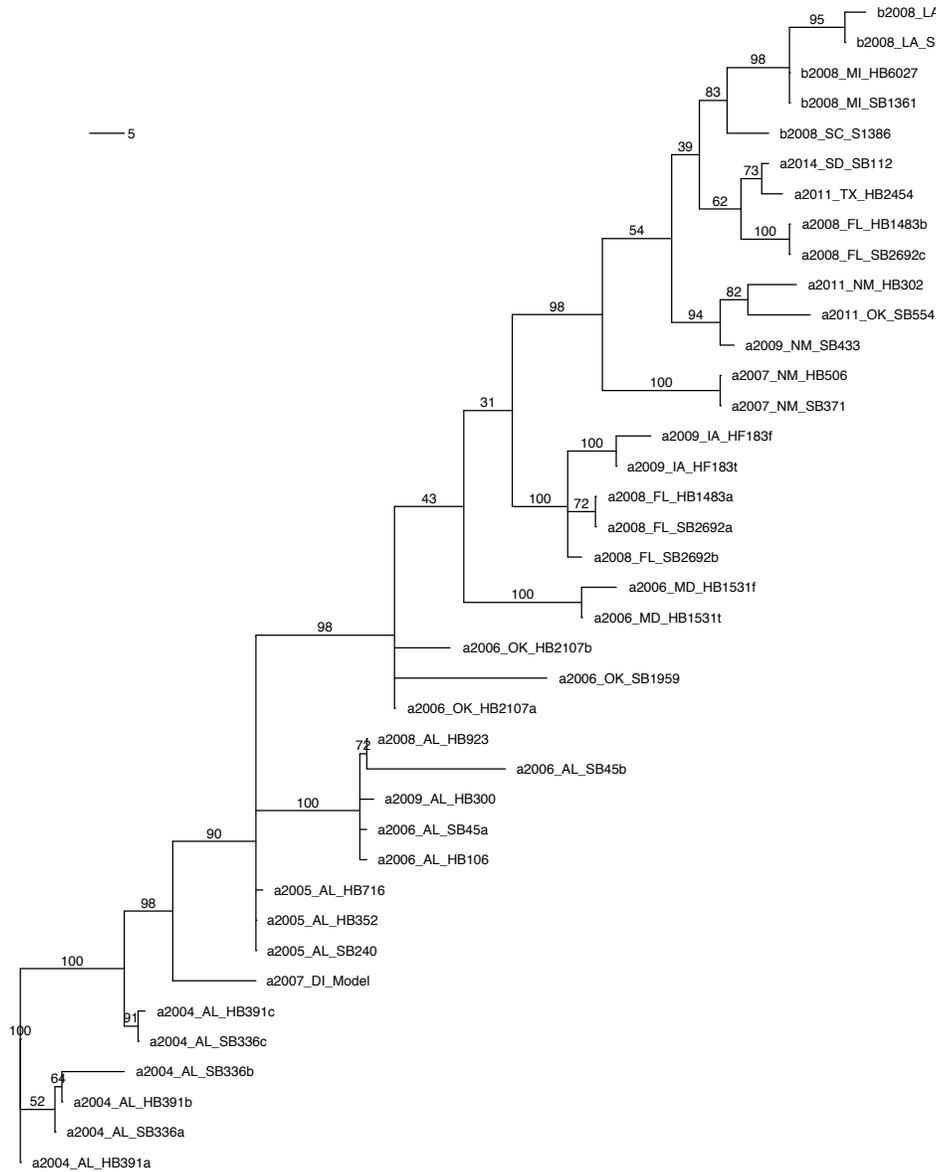


Fig. S5. MP strict consensus tree inferred from the “group a”, AFA OTUs (with the LSEA and a few “group b” SEA relatives added to see where the SEA tradition connects). The tree exhibits high congruence with OTU date. Branch lengths represent the minimum number of character steps. Numbers are the bootstrap support out of 100 bootstrap trees.

**MP tree
(bootstrap support/frequency)**

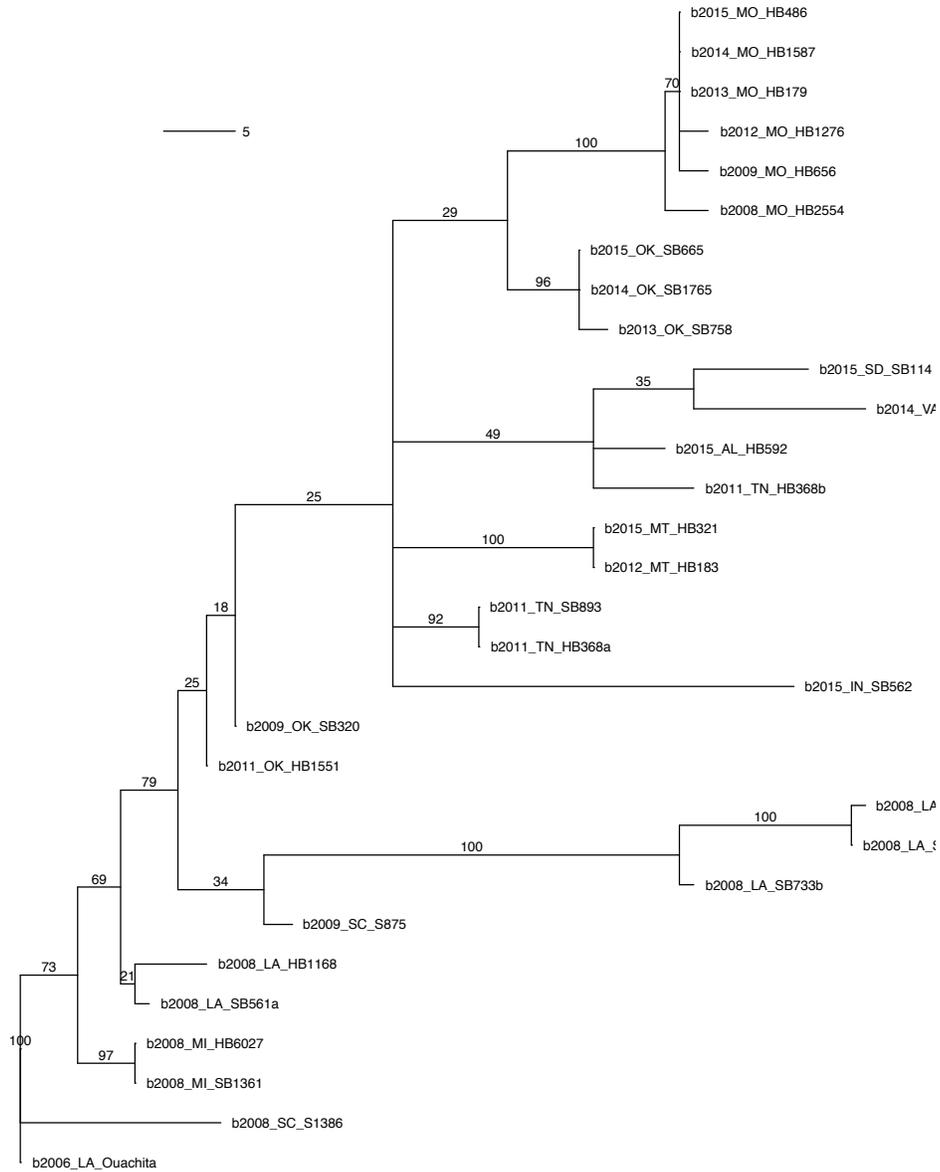


Fig. S6. MP strict consensus tree inferred from the “group b”, SEA OTUs. The tree exhibits high congruence with OTU date when rooted on the 2006 Ouachita bill, reflecting the known fact that Ouachita was the origin of much of the SEA language. Branch lengths represent the minimum number of character steps. Numbers are the bootstrap support out of 100 bootstrap trees; these are somewhat lower than for the AFA tree, due to fewer characters.

**MP tree
(bootstrap support/frequency)**

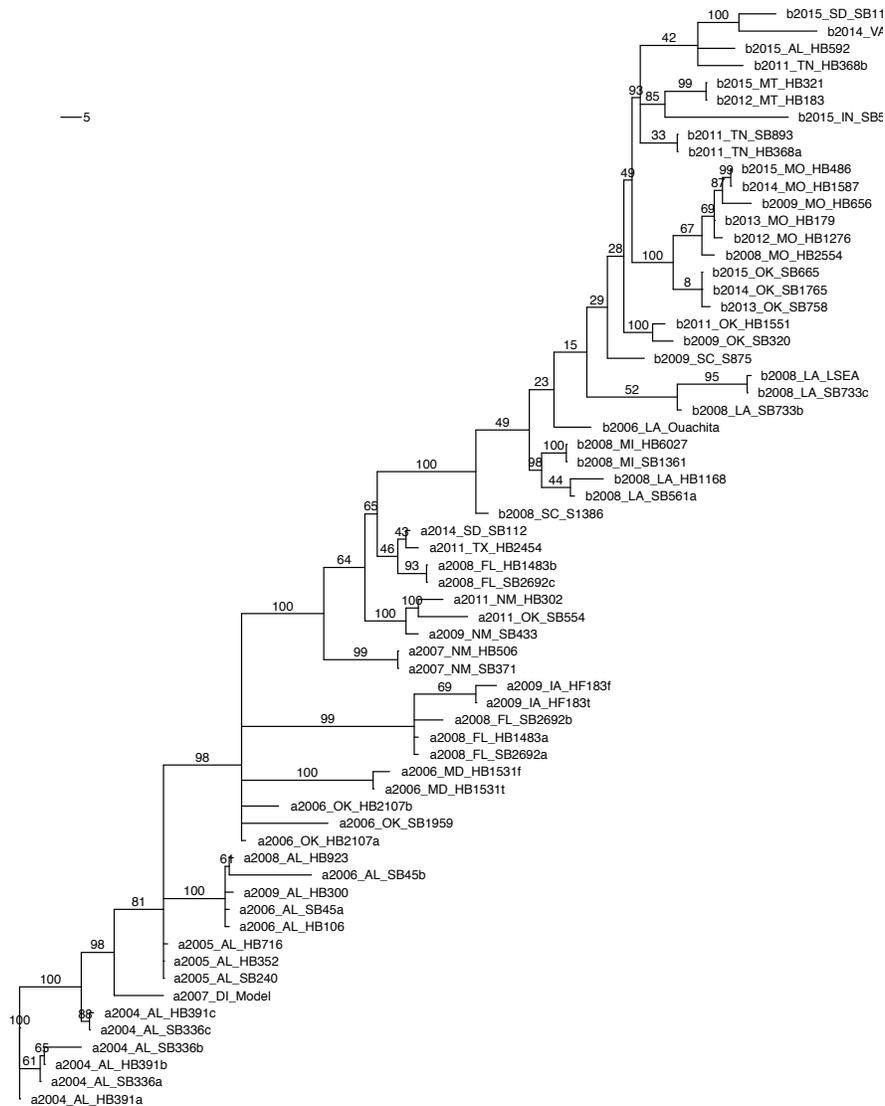


Fig. S7. MP strict consensus tree inferred from all characters and all taxa, with no constraints. The low bootstrap values near the Ouachita OTU, and the observed temporal incongruence (some 2008 bills before the 2006 Ouachita known to be the source of some of their text) indicate that the root of the SEA subtree is being misplaced due to the absence of AFA characters in the Ouachita OTU. The SEA subtree was constrained to be rooted on Ouachita in subsequent analysis, with other topology left free. Branch lengths represent minimum number of character changes. Numbers are the bootstrap support out of 100 bootstrap trees.

**Dated phylogeny of antievolution bills
(Beast2 MCC summary, sampled ancestor BD tree prior)**

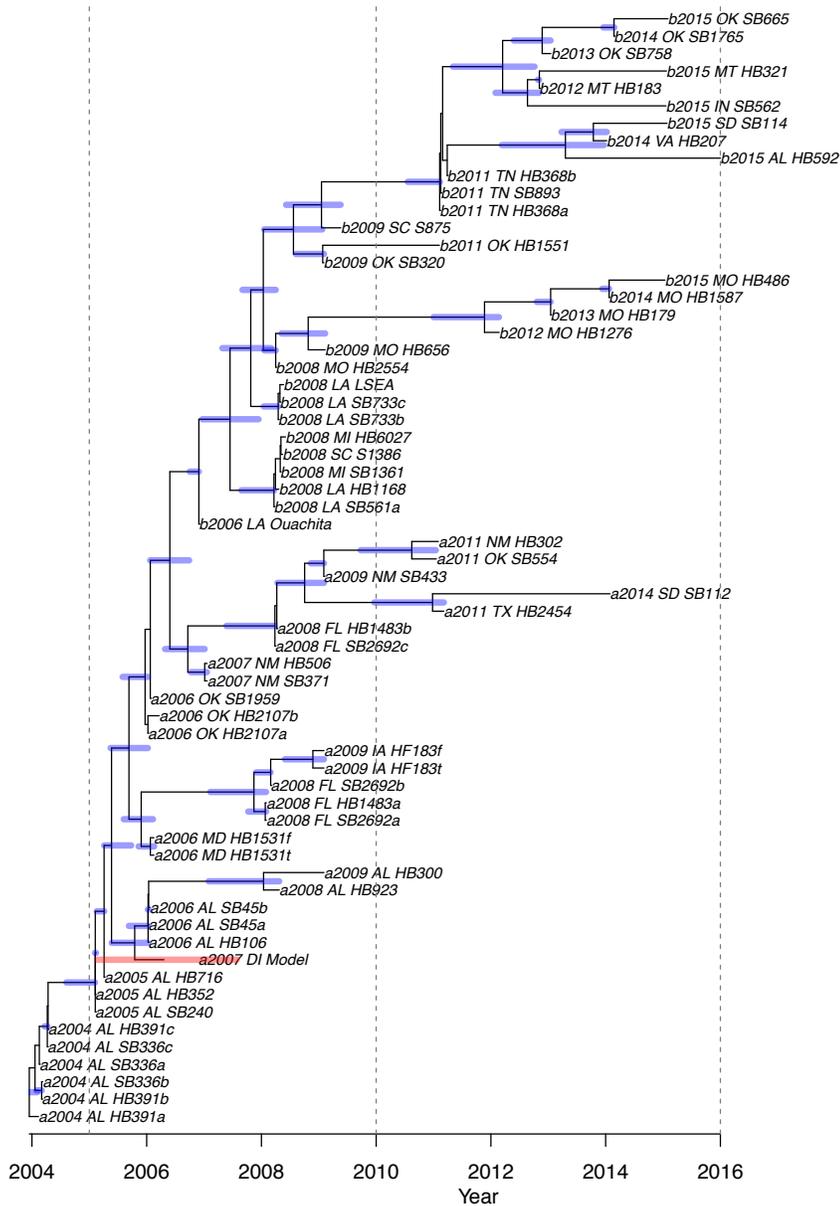


Fig. S9. Bayesian tip-dated phylogeny of the antievolution policies. Posterior probabilities (PP) of each bipartition are not shown as many branches are very short, but most branches have high support (only 6 branches with $PP < 0.5$, 40 branches with $PP > 0.85$). Blue bars represent the 95% highest posterior density (HPD) of each node date. HPD widths are typically < 1 year. The red bar represents the 95% HPD for the tip date of the DI model legislation, when given a flat prior distribution of $Uniform(1/1/2004, 9/7/2007)$. For the posterior density of the date of the DI Model Bill, see Fig. S10B. The maximum posterior probability estimate for the tip date is early 2006, suggesting a close kinship to the 2005 and 2006 Alabama legislative proposals.

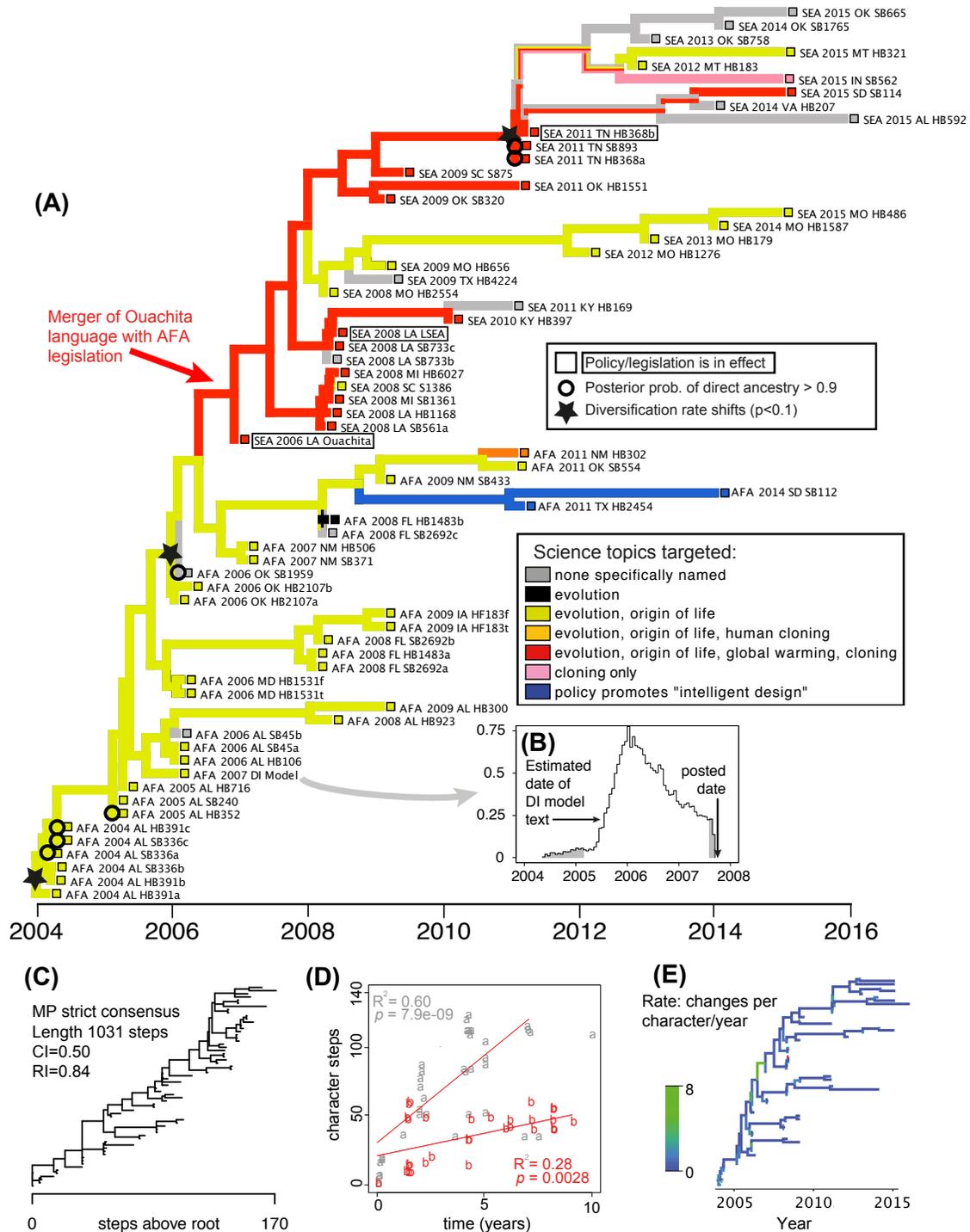


Fig. S10. Tracing the evolution of antievolution legislation. (A) Maximum clade credibility tree from Bayesian tip-dating analysis of 67 policies. The Science Education Acts (SEAs) originated by combining text from the Academic Freedom Acts (AFAs) with policy text from 2006 Ouachita Parish, Louisiana. Branch colors indicate the sciences targeted; mixed colors on a branch indicate uncertainty in the reconstruction. Seven bills have a high posterior probability of being direct ancestors of the rest of the tradition (circles). The nodes showing the strongest evidence for shifts in diversification

rate (stars) tend to be close to these direct ancestors. Tip labels indicate AFA/SEA, year, state, bill number (SB/HB=Senate/House Bill), and finally versions (a/b/c=legislative revisions, t/f=teachers or faculty targeted). (B) Estimating the date of the Discovery Institute model bill. Figs. C-E validate the assumptions of the Bayesian analysis. (C) Undated tree where branch lengths are in number of reconstructed changes, estimated with parsimony. (D) Testing the Bayesian model's "clock" assumption by regressing tip heights from C on tip dates from A: AFAs (grey, a) and SEAs (red, b) both show significant correlation; the flatter slope of the SEAs is primarily due to fewer characters. (E) Average rate of change per character per year from the Bayesian analysis. The spike in rates in the middle of the tree is driven by the addition of Ouachita-derived characters.

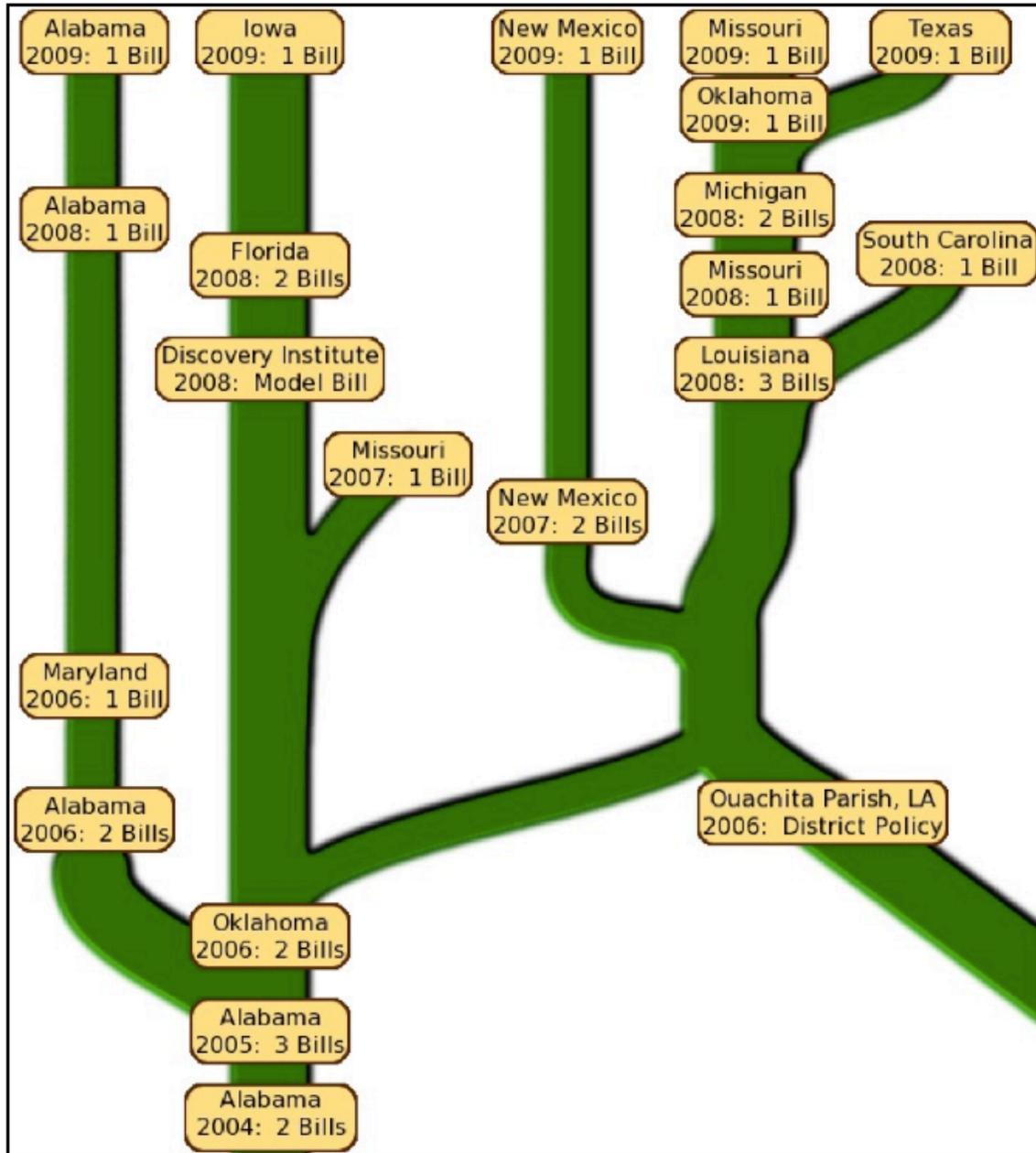


Fig. S11. Manually-generated phylogeny of antievolution legislation produced at NCSE in 2010. Created by Anton Mates. Copyright 2010 NCSE, reproduced with permission.

Additional Data S1 (separate file)

The text alignment, NEXUS data files, and code files are available as a zipfile in Supplemental Data S1.