

Empirical calibrated radiocarbon sampler: a tool for incorporating radiocarbon-date and calibration error into Bayesian phylogenetic analyses of ancient DNA

MARTYNA MOLAK,*† MARC A. SUCHARD,‡§¶ SIMON Y. W. HO,* DAVID W. BEILMAN** and BETH SHAPIRO††

*School of Biological Sciences, University of Sydney, Edgeworth David building A11, Science Road, 2006, Sydney, NSW, Australia, †Museum and Institute of Zoology, Polish Academy of Sciences, Wilcza 64, 00-679, Warsaw, Poland, ‡Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles 695 Charles E. Young Drive South, Los Angeles, CA, 90095, USA, §Department of Biostatistics, UCLA Fielding School of Public Health, University of California, 650 Charles E. Young Drive South, Los Angeles, CA, 90095, USA, ¶Department of Biomathematics, David Geffen School of Medicine at UCLA, University of California, 621 Charles E. Young Drive South, Los Angeles, CA, 90095-1766, USA, **Department of Geography, University of Hawaii, 445 Saunders Hall, 2424 Maile Way, Honolulu, HI, 96822, USA, ††Department of Ecology and Evolutionary Biology, University of California, 1156 High Street, Santa Cruz, CA, 95064, USA

Abstract

Studies of DNA from ancient samples provide a valuable opportunity to gain insight into past evolutionary and demographic processes. Bayesian phylogenetic methods can estimate evolutionary rates and timescales from ancient DNA sequences, with the ages of the samples acting as calibrations for the molecular clock. Sample ages are often estimated using radiocarbon dating, but the associated measurement error is rarely taken into account. In addition, the total uncertainty quantified by converting radiocarbon dates to calendar dates is typically ignored. Here, we present a tool for incorporating both of these sources of uncertainty into Bayesian phylogenetic analyses of ancient DNA. This empirical calibrated radiocarbon sampler (ECRS) integrates the age uncertainty for each ancient sequence over the calibrated probability density function estimated for its radiocarbon date and associated error. We use the ECRS to analyse three ancient DNA data sets. Accounting for radiocarbon-dating and calibration error appeared to have little impact on estimates of evolutionary rates and related parameters for these data sets. However, analyses of other data sets, particularly those with few or only very old radiocarbon dates, might be more sensitive to using artificially precise sample ages and should benefit from use of the ECRS.

Keywords: age-estimation error, ancient DNA, BEAST, phylogenetic dating

Received 2 April 2014; revision received 6 June 2014; accepted 9 June 2014

Introduction

Genetic analysis of ancient and historical samples provides access to valuable information that is not available from modern samples alone. For example, phylogenetic and phylogeographic analyses incorporating serially-sampled data have allowed estimates of the relationships between extinct species (e.g. Bunce *et al.* 2003), inference of past population dynamics (e.g. Lorenzen *et al.* 2011), and insights into hominin evolution (e.g. Reich *et al.* 2011; Fu *et al.* 2013). Ancient DNA data can also be used to estimate evolutionary rates and associated timescales, using the ages of the ancient samples to calibrate the

molecular clock (Rambaut 2000; Drummond *et al.* 2003). Molecular-clock analyses of ancient DNA have been particularly informative about evolutionary rates over population timescales (Ho *et al.* 2011), providing estimates, for example, of the timing of migration events (Debruyne *et al.* 2008; Edwards *et al.* 2011).

With the exception of historical samples that have documented dates of collection, the ages of ancient samples are typically unknown and need to be estimated. Radiocarbon dating (dating using decay of ^{14}C), by scintillation counting or by accelerator mass spectrometry (AMS), is a common method for estimating sample ages and has a theoretical and methodological foundation that provides a quantifiable amount of uncertainty that can be rather considerable (Guilderson *et al.* 2005). In phylogenetic analyses of ancient DNA, sample ages are typi-

Correspondence: Martyna Molak, Fax: +48 226296302; E-mail: martyna.molak@gmail.com

cally assigned the mean or median single value of the age distribution and the rest of the uncertainty information is ignored. Methods were recently implemented in the software package *BEAST* (Drummond *et al.* 2012) to allow radiocarbon dating or other sources of error to be taken into account, by specifying a prior distribution on the age of each sample (Ho & Phillips 2009; Shapiro *et al.* 2011). This approach has been used to incorporate uncertainty when direct AMS radiocarbon dates are not available, for example where ages are inferred from stratigraphic information (e.g. Orlando *et al.* 2013; Stiller *et al.* In press).

Previous work has suggested that incorporating error associated with AMS radiocarbon age estimates tends to have a limited impact on estimates of evolutionary and demographic parameters (Molak *et al.* 2013). However, there might be instances in which this error plays an important role in the analysis. For example, if the estimated error is large (as it tends to be for many samples towards the upper limit of ca. 40–50 000 years for ^{14}C dating) or when only one or a few ancient sequences are used, ignoring the error could lead to artificially precise estimates of the evolutionary rate. As a consequence, estimates of the timing of demographic events would be misleadingly precise.

Moreover, radiocarbon ages determined from ^{14}C values and the accepted radioisotope half-life differ from absolute (calendar) ages because the atmospheric concentration of ^{14}C has varied through time. If calendar ages are desired, then the radiocarbon ages need to be converted using a calibration curve. Calibration curves are based on analysis of growth patterns correlated with calendar years, such as those observed in tree rings or coral, and comparison of these with their radiocarbon ages. Obtaining a calibration curve for the entire age range spanned by radiocarbon-dating methods requires the combination of several sources of calibration, and curves continue to improve as more data become available and methodology improves (Stuiver & Reimer 1993; Reimer *et al.* 2013). Importantly, the uncertainty quantified by converting radiocarbon years before present (^{14}C yBP) to calendar years before present (cal yBP) is compounded with that of the initial ^{14}C measurement error (Fig. 1). Probability distributions of calibrated ages usually do not follow a simple parametric distribution and are often multimodal, making it a challenge to incorporate this uncertainty into a phylogenetic analysis.

Here, we present the empirical calibrated radiocarbon sampler (ECSR), which we have implemented in the Bayesian phylogenetic software *BEAST* v1.8 (Drummond *et al.* 2012). The ECSR allows the uncertainty in calibrated radiocarbon dates to be taken into account directly as nonparametric prior information. This is achieved by providing the software with an empirical

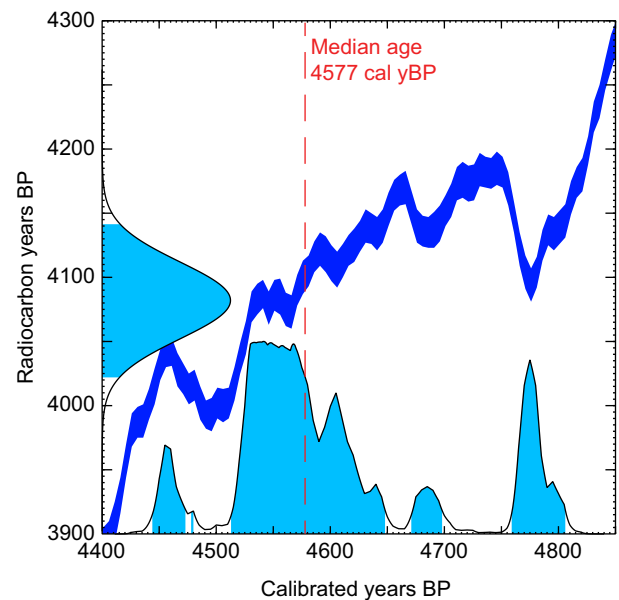


Fig. 1 Example of a calibration plot with a multimodal probability distribution. The radiocarbon age estimate of the sample is 4082 ± 30 ^{14}C yBP. The median calibrated age estimate of the sample is 4577 cal yBP. The probability density functions of the uncalibrated and calibrated ages are plotted along the *y*- and *x*-axes, respectively. Conversion between the two is performed using IntCal13 calibration curve (in dark blue).

description of the probability density function for the calibrated age estimate for each radiocarbon-dated sample. The age of that sample is then integrated over this probability distribution through a Markov chain Monte Carlo (MCMC) algorithm. We introduce a simple method to incorporate age uncertainty into Bayesian phylogenetic analysis using the probability density functions generated through the calibration program *CALIB* (Stuiver & Reimer 1993).

Methods

BEAST analysis with the ECSR

We implemented in *BEAST* v1.8 a novel nonparametric probability distribution to model empirical information about uncertainty and systematic error in specifying tip ages. The ECSR takes as input age probability density files generated by *CALIB* 7.0.0 (or newer) that provide a finite grid of ages and their associated probability masses (alternative software can be used to generate equivalent probability density files). The novel *BEAST* implementation reads in these values, assumes a 1st-order spline fit between these grid values to allow for continuous ages, places zero density outside the minimal and maximal grid values and renormalizes to generate an integrable density function for use with MCMC.

In making this transformation for our examples in this study, the ECRS applies an offset of 60 years to the age values in the probability density text files. Data reporting conventions for ^{14}C dating are such that ages are given in 'years before 1950'. Therefore, the ECRS assumes that modern (age = 0) samples were collected in 2010. For evolutionary rate estimation, any disparity between 2010 and the real age of the modern samples used should be negligible. This offset was introduced in the ECRS to avoid age distributions of young radiocarbon-dated samples extending to negative values.

The age distributions cannot be easily applied to data sets that include no modern samples (e.g. extinct species). When there is no stable zero-height point, the current implementation of BEAST will reassign the timescale of the tree according to the youngest sequence sampled in any given MCMC step. This problem can be avoided by fixing the age of the youngest sample in the data set to a point value, provided that none of the other sample-age distributions can cross this point. Otherwise, they too will cause the tree to be rescaled. In such circumstances, we advise fixing the ages of all the sequences of which age distributions cross the value of the age of the youngest sample.

The BEAST input file needs to be edited to define the age prior for each ancient DNA sample as the probability density defined in the age probability density text file. Instructions on how to use the ECRS, along with all used XML files, are available as Supporting Information (Data S1–S6).

Testing the ECRS

We tested the ECRS on three published ancient DNA data sets: bison (Shapiro *et al.* 2004), muskox (Campos *et al.* 2010) and human (Fu *et al.* 2013). Sequence alignments of the mitochondrial control region were used for bison (591 bp, 159 sequences) and muskox (682 bp, 131 sequences), whereas whole mitochondrial genomes were used for humans (16 564 bp, 64 sequences). Ages of ancient DNA sequences ranged from 125 to 43 400 ^{14}C yBP for bison (142 samples), 115–45 900 ^{14}C yBP for muskox (117 samples) and 690–39 475 ^{14}C yBP for humans (10 samples).

We performed a Bayesian phylogenetic analysis of each data set, using either the median calibrated ages of the samples or with the ECRS using the age distributions obtained with the IntCal13 calibration curve (Reimer *et al.* 2013). The probability density for the calibrated age of each radiocarbon-dated specimen was obtained using CALIB 7.0.0 (Stuiver & Reimer 1993). Because the ECRS adds 60 years to the age of each ancient DNA sequence, we added 60 years also to the median calibrated ages of the ancient DNA sequences for ECRS testing. This

allowed us to compare directly the analyses using median calibrated sample ages and using the age probability densities.

For each data set, the best-fitting model of nucleotide substitution was chosen according to the Bayesian Information Criterion using ModelGenerator (Keane *et al.* 2006). Each analysis was performed in BEAST v1.8 using a conditional reference prior (Ferreira & Suchard 2008) for the substitution rate. The sample ages were the only source of calibrating information for the molecular clock. For each data set, we performed analyses using both a constant-size coalescent model and the skygrid model (Gill *et al.* 2013). Owing to the intraspecific nature of the data, we assumed a strict molecular clock. Markov chains were run for at least 10^7 steps, with parameters subsampled every 10^3 steps. Chains were run for longer if required until all parameter estimates converged and achieved effective sample sizes of >100 . In each analysis, the first 10% of steps were discarded as burn-in. For each data set, we compared marginal likelihoods estimated for both coalescent models using stepping-stone and path sampling protocols (Baele *et al.* 2012, 2013). This Bayes-factor analysis supported the Bayesian skygrid model for bison and human, and a constant-size model for muskox.

We used Bayes factors (estimated as above) to compare the analyses calibrated using median sample ages and those based on the ECRS (Suchard *et al.* 2001). We also investigated estimates of substitution rates and the age of the root, as well as the duration of the analysis.

Results of ECRS testing

The ECRS, as implemented in the BEAST phylogenetic framework, makes it possible to account for the uncertainty in radiocarbon measurement and age calibration when estimating evolutionary rates. The model performs as expected, as is reflected in the posterior distributions of sample ages, which accurately recapitulate the probability density associated with the calibrated age of each specimen (Fig. 2).

Accounting for this uncertainty in the sample ages did not lead to a substantial improvement in marginal likelihood for any of the three data sets analysed (Table 1). The 95% credibility intervals for the substitution rate and age of the root were of similar widths to or slightly narrower (up to 8%) than the ones obtained using median sample ages for molecular-clock calibration. We also note that, as expected, the computation time (per effective sample size of 100) was longer for the ECRS model than for the model that used median sample age for two out of the three data sets (bison and muskox). The amount of computation time will depend on

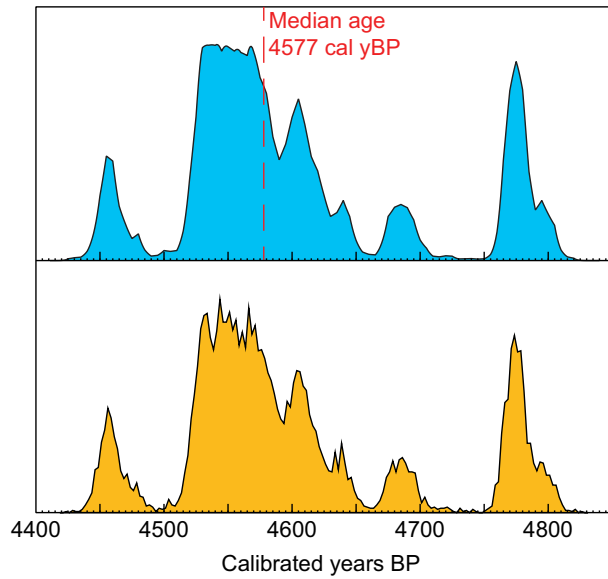


Fig. 2 Example of a posterior distribution of a sample age (bottom plot, orange) compared with the calibration plot obtained in CALIB (top plot, blue) and used as a prior for Bayesian phylogenetic analysis. The vertical axis represents the probability density. Note that the figure shows ages in yBP (before 1950), whereas empirical calibrated radiocarbon sampler (E CRS) yields posterior ages in years before 2010.

specific characteristics of the data sets, such as alignment length and number of sequences and on the availability of resources.

Discussion

We have presented the E CRS, a tool implemented in the BEAST software package that enables the uncertainty from radiocarbon measurement and age calibration to be taken into account in Bayesian phylogenetic analysis. Concordance between the prior and posterior distributions of sample ages indicates that there is no conflict between the prior age distribution and the genetic signal in the data set. However, it also indicates that the genetic signal is not strong enough in any of our three data sets to improve the precision of the age estimate, even for the human data set which comprises complete mitochondrial genomes.

Accounting for the uncertainty in sample ages in Bayesian phylogenetic analysis appears to have had a limited impact on estimates of key evolutionary parameters, including the evolutionary rate and the age of the root of the tree. This result is consistent with previous findings from a range of ancient DNA data (Molak *et al.* 2013), and suggests that the common practice of ignoring errors in sample-age estimation and radiocarbon calibration in phylogenetic analyses of ancient DNA probably

Table 1 Comparison of Bayesian phylogenetic analyses made using the empirical calibrated radiocarbon sampler (E CRS) and median sample ages. Data shown include estimates and 95% highest posterior density (HPD) intervals for the substitution rate and age of the root; speed of the BEAST run; and Bayes factors for the E CRS compared with median sample ages, with marginal likelihoods calculated using stepping-stone and path sampling

	Population model	Substitution model	Sample ages used	Substitutions/site/year			Time to most recent common ancestor			Run speed		Bayes factors	
				Median estimate	95% HPD	95% HPD width	Median estimate	95% HPD width	95% HPD width ratio*	Run speed ratio†	Via stepping-stone sampling*	Via path sampling*	
Bison	Skygrid	TrN+G ₆	Median E CRS	4.39×10^{-7}	3.26×10^{-7} – 5.60×10^{-7}	2.34×10^{-7}	1.08×10^5	8.81×10^4 – 1.37×10^5	4.86×10^4	1.03	2.47	0.00	0.00
			Median E CRS	4.39×10^{-7}	3.25×10^{-7} – 5.60×10^{-7}	2.35×10^{-7}	1.08×10^5	8.86×10^4 – 1.37×10^5	5.03×10^4	0.92	0.62	1.63	2.22
Human	Skygrid	TrN+G ₆	Median E CRS	1.99×10^{-8}	1.58×10^{-8} – 2.42×10^{-8}	8.37×10^{-9}	1.54×10^5	1.22×10^5 – 1.96×10^5	7.36×10^4	0.92	0.62	1.63	2.22
			Median E CRS	2.03×10^{-8}	1.64×10^{-8} – 2.43×10^{-8}	7.90×10^{-9}	1.53×10^5	1.22×10^5 – 1.90×10^5	6.74×10^4	1.00	4.45	0.91	0.10
Musk ox	Constant	TVM+G ₆	Median E CRS	8.16×10^{-7}	6.20×10^{-7} – 1.02×10^{-6}	3.99×10^{-7}	8.04×10^4	6.78×10^4 – 9.51×10^4	2.72×10^4	1.00	4.45	0.91	0.10
			Median E CRS	8.09×10^{-7}	6.23×10^{-7} – 9.99×10^{-7}	3.76×10^{-7}	8.05×10^4	6.74×10^4 – 9.47×10^4	2.72×10^4	1.00	4.45	0.91	0.10

*E CRS vs median age.

†E CRS vs median age run speed calculated as hours per 100 ESS for rate estimate.

does not introduce significant error into these analyses. Nevertheless, implementing the ECRS is an important step towards more accurate models of biological parameters for phylogenetic analysis.

The ECRS is simple to implement and reduces the potential effect of disregarding sample-age-estimation error in Bayesian phylogenetic analyses. If the sequence ages used to calibrate the molecular clock are artificially precise, our confidence in phylogenetic estimates of rates and timescales could be overstated. Consequently, accounting for radiocarbon-dating and calibration error might be most useful in analyses of data sets that include only a small number of ancient DNA sequences, where each sample age constitutes a considerable proportion of the overall tree-calibrating information. This effect will be particularly pronounced when the sequences have ages for which calibration curves produce large age-estimation errors, as is typically the case for samples that are near the upper age limit for radiocarbon dating.

Acknowledgements

MM is supported by a University of Sydney International Scholarship. MAS is supported in part by National Institutes of Health grants R01 AI107034 and R01 HG006139 and National Science Foundation grants DMS 1264153 and IIS 1251151. SYWH is supported by the Australian Research Council. DWB is supported in part by National Science Foundation grants 1108116 and 1246359. BS is supported by The Packard Foundation and National Science Foundation grant 1203990.

References

- Baele G, Lemey P, Bedford T *et al.* (2012) Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution*, **29**, 2157–2167.
- Baele G, Li WLS, Drummond AJ, Suchard MA, Lemey P (2013) Accurate model selection of relaxed molecular clocks in Bayesian phylogenetics. *Molecular Biology and Evolution*, **30**, 239–243.
- Bunce M, Worthy TH, Ford T *et al.* (2003) Extreme reversed sexual size dimorphism in the extinct New Zealand moa *Dinornis*. *Nature*, **425**, 172–175.
- Campos PF, Willerslev E, Sher A *et al.* (2010) Ancient DNA analyses exclude humans as the driving force behind late Pleistocene musk ox (*Ovibos moschatus*) population dynamics. *Proceedings of the National Academy of Sciences, USA*, **107**, 5675–5680.
- Debruyne R, Chu G, King CE *et al.* (2008) Out of America: ancient DNA evidence for a new world origin of late quaternary woolly mammoths. *Current Biology*, **18**, 1320–1326.
- Drummond AJ, Pybus OG, Rambaut A, Forsberg R, Rodrigo AG (2003) Measurably evolving populations. *Trends in Ecology and Evolution*, **18**, 481–488.
- Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.
- Edwards CJ, Suchard MA, Lemey P *et al.* (2011) Ancient hybridization and an Irish origin for the modern polar bear matriline. *Current Biology*, **21**, 1251–1258.

- Ferreira MAR, Suchard MA (2008) Bayesian analysis of elapsed times in continuous-time Markov chains. *Canadian Journal of Statistics*, **36**, 355–368.
- Fu Q, Mitnik A, Johnson PL *et al.* (2013) A revised timescale for human evolution based on ancient mitochondrial genomes. *Current Biology*, **23**, 553–559.
- Gill MS, Lemey P, Faria NR *et al.* (2013) Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. *Molecular Biology and Evolution*, **30**, 713–724.
- Guilderson TP, Reimer PJ, Brown TA (2005) The boon and bane of radiocarbon dating. *Science*, **307**, 362–364.
- Ho SYW, Phillips MJ (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, **58**, 367–380.
- Ho SYW, Lanfear R, Bromham L *et al.* (2011) Time-dependent rates of molecular evolution. *Molecular Ecology*, **20**, 3087–3101.
- Keane TM, Creevey CJ, Pentony MM, Naughton TJ, McLnerney JO (2006) Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evolutionary Biology*, **6**, 29.
- Lorenzen ED, Nogués-Bravo D, Orlando L *et al.* (2011) Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature*, **479**, 359–364.
- Molak M, Lorenzen ED, Shapiro B, Ho SYW (2013) Phylogenetic estimation of timescales using ancient DNA: the effects of temporal sampling scheme and uncertainty in sample ages. *Molecular Biology and Evolution*, **30**, 253–262.
- Orlando L, Ginolhac A, Zhang G *et al.* (2013) Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature*, **499**, 74–78.
- Rambaut A (2000) Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. *Bioinformatics*, **16**, 395–399.
- Reich D, Patterson N, Kircher M *et al.* (2011) Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania. *American Journal of Human Genetics*, **89**, 516–528.
- Reimer PJ, Bard E, Bayliss A *et al.* (2013) IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP. *Radiocarbon*, **55**, 1869–1887.
- Shapiro B, Drummond AJ, Rambaut A *et al.* (2004) Rise and fall of the Beringian steppe bison. *Science*, **306**, 1561–1565.
- Shapiro B, Ho SYW, Drummond AJ *et al.* (2011) A Bayesian phylogenetic method to estimate unknown sequence ages. *Molecular Biology and Evolution*, **28**, 879–887.
- Stiller M, Molak M, Prost S *et al.* (In press) Mitochondrial DNA diversity and evolution of the Pleistocene cave bear complex. *Quaternary International*, doi: 10.1016/j.quaint.2013.09.023
- Stuiver M, Reimer PJ (1993) Extended ¹⁴C data base and revised CALIB 3.0 ¹⁴C age calibration program. *Radiocarbon*, **35**, 215–230.
- Suchard MA, Weiss RE, Sinsheimer JS (2001) Bayesian selection of continuous-time Markov chain evolutionary models. *Molecular Biology and Evolution*, **18**, 1001–1013.

M.M. planned and performed the analyses, wrote the manuscript. M.A.S. developed the new empirical distribution, implemented this distribution in BEAST, planned the analyses and wrote the manuscript. S.Y.W.H. planned the analyses and wrote the manuscript. D.W.B. provided calibration data and wrote the manuscript. B.S. planned and performed the analyses and wrote the manuscript.

Data Accessibility

All the genetic and radiocarbon-dating data used in the analyses used have been previously published and are available alongside the original publications (with DNA sequences available via GenBank): bison – Shapiro *et al.* 2004; human – Fu *et al.* 2013; muskox – Campos *et al.* 2010. For the ECRS testing, we chose from these data sets only the samples for which ^{14}C ages and errors were provided and of which ages did not exceed the limits of the IntCal13 calibration curve. All input XML files used to test the ECRS (including the script required to run BEAST with ECRS) and information about the ages of the samples (provided as CALIB input files), which will allow the readers to fully replicate our analyses is available as Supporting Information.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1 CALIB CSV input file and BEAST XML input files for analyses with constant-size coalescent model and skygrid model calibrated using median sample ages or ECRS for the bison data set (Shapiro *et al.* 2004) used for the ECRS testing.

Data S2 CALIB CSV input file and BEAST XML input files for analyses with constant-size coalescent model and skygrid model calibrated using median sample ages or ECRS for the human data set (Fu *et al.* 2013) used for the ECRS testing.

Data S3 CALIB CSV input file and BEAST XML input files for analyses with constant-size coalescent model and skygrid model calibrated using median sample ages or ECRS for the muskox data set (Campos *et al.* 2010) used for the ECRS testing.

Data S4 An example file with blocks of script generated for pasting into the BEAST XML input file to implement the ECRS (human data set; Fu *et al.* 2013).

Data S5 ECRS XML script blocks template - a file generating blocks of script that can be pasted into the indicated sections of the BEAST XML input file to enable data analysis using the ECRS.

Data S6 Instructions for use of the ECRS.