letters to nature

consistent with that observed[28], and HadCM3, when it includes both anthropogenic and natural forcings, simulates many features of observed twentieth-century temperature change[22], indicating some success in incorporating external forcings including those due to solar changes and volcanic aerosol. Nevertheless, the current dependence on model-based rather than observationally based estimates of natural variability needs to be tested further against observational evidence. We do not include the effect of observational error in our analysis. The effect of observational sampling error on detection and attribution results has been shown to be small[29], but we do not as yet have an estimate of the effects of systematic instrumental errors, such as changes in measurement practices or urbanization.

1. Forest, C. E., Stone, P. H., Sokolov, A. P., Allen, M. R. & Webster, M. Quantifying uncertainties in climate system properties using recent climate observations. *Science* **295**, 113–117 (2002).
2. Knutti, R., Stocker, T. F., Joos, F. & Plattner, G.-K. Constraints on radiative forcing and future climate change from observations and climate model ensembles. *Nature* **416**, 719–723 (2002).
3. Allen, M. R., Raper, S. C. B. & Mitchell, J. F. B. Uncertainty in the IPCC's third assessment report. *Science* **293**, 430–433 (2001).
4. Allen, M. R., Stott, P. A., Mitchell, J. F. B., Schnur, R. & Delworth, T. L. Uncertainty in forecasts of anthropogenic climate change. *Nature* **407**, 617–620 (2000).
5. Nakicenovic, N. & Swart, R. *Special Report on Emission Scenarios* (Cambridge Univ. Press, Cambridge, 2000).
6. Wigley, T. M. L. & Raper, S. C. B. Interpretation of high projections for global-mean warming. *Science* **293**, 451–454 (2001).
7. Reilly, J. *et al.* Uncertainty and climate change assessments. *Science* **293**, 430–433 (2001).
8. Allen, M. R. & Tett, S. F. B. Checking for model consistency in optimal fingerprinting. *Clim. Dyn.* **15**, 419–434 (1999).
9. Tett, S. F. B., Stott, P. A., Allen, M. R., Ingram, W. J. & Mitchell, J. F. B. Causes of twentieth century temperature change near the Earth's surface. *Nature* **399**, 569–572 (1999).
10. Cubasch, U. *et al.* in *Climate Change 2001: The Scientific Basis* Ch. 9 (eds Houghton, J. T. *et al.*) 525–582 (Cambridge Univ. Press, Cambridge, 2001).
11. Forest, C. E., Allen, M. R., Stone, P. H. & Sokolov, A. P. Constraining uncertainties in climate models using climate change detection techniques. *Geophys. Res. Lett.* **27**, 569–572 (2000).
12. Booth, B. B. B., Allen, M. R. & Kettleborough, J. A. *Exploring the Linearity of the Climate Response to External Forcing* (Tech. Report RAL-TR2002-005, Rutherford Appleton Laboratory, Chilton, UK, 2002).
13. Houghton, J. T. *et al.* (eds) *Climate Change 2001: The Scientific Basis* (Cambridge Univ. Press, Cambridge, 2001).
14. Stocker, T. F. & Schmittner, A. Influence of CO_2 emission rates on the stability of the thermohaline circulation. *Nature* **388**, 862–865 (1997).
15. Cox, P. N., Betts, R. A., Jones, C. D., Spall, S. A. & Totterdell, I. J. Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature* **408**, 184–187 (2000).
16. Gordon, C. *et al.* The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Clim. Dyn.* **16**, 147–168 (2000).
17. Pope, V. D., Gallani, M. L., Rowntree, P. R. & Stratton, R. A. The impact of new physical parametrizations in the Hadley Centre climate model—HadAM3. *Clim. Dyn.* **16**, 123–146 (2000).
18. Twomey, S. A. Pollution and the planetary albedo. *Atmos. Environ.* **8**, 1251–1256 (1974).
19. Sato, M., Hansen, J. E., McCormick, M. P. & Pollack, J. B. Stratospheric aerosol optical depths (1850–1990). *J. Geophys. Res.* **98**, 22987–22994 (1993).
20. Lean, J., Beer, J. & Bradley, R. Reconstruction of solar irradiance since 1610: implications for climate change. *Geophys. Res. Lett.* **22**, 3195–3198 (1995).
21. Johns, T. C. *et al. Anthropogenic Climate Change for 1860 to 2100 Simulated with the HadCM3 Model under Updated Emissions Scenarios* (Tech. Note 22, Hadley Centre for Climate Prediction and Research, Bracknell, UK, 2001).
22. Stott, P. A. *et al.* External control of 20th century temperature by natural and anthropogenic forcings. *Science* **290**, 2133–2137 (2000).
23. Allen, M. R. Do-it-yourself climate prediction. *Nature* **401**, 642 (1999).
24. Stott, P. A. *et al.* Attribution of twentieth century temperature change to natural and anthropogenic causes. *Clim. Dyn.* **17**, 1–21 (2001).
25. Stott, P. A. & Tett, S. F. B. Scale-dependent detection of climate change. *J. Clim.* **11**, 32892–3294 (1998).
26. Hasselmann, K. Multi-pattern fingerprint method for detection and attribution of climate change. *Clim. Dyn.* **13**, 601–611 (1997).
27. Allen, M. R. & Stott, P. A. Estimating signal amplitudes in optimal fingerprinting. Part I: theory. *Clim. Dyn.* (submitted).
28. Crowley, T. J. Causes of climate change over the past 1000 years. *Science* **289**, 270–277 (2000).
29. Hegerl, G. C., Jones, P. D. & Barnett, T. P. Effect of observational sampling error on the detection and attribution of anthropogenic climate change. *J. Clim.* **14**, 198–207 (2001).

Using the fossil record to estimate the age of the last common ancestor of extant primates

Simon Tavaré*, **Charles R. Marshall**†, **Oliver Will**‡, **Christophe Soligo**§‖ **& Robert D. Martin**¶

* *Departments of Biological Sciences, Mathematics, and Preventive Medicine, University of Southern California, Los Angeles, California 90089-1340, USA*
† *Departments of Organismic and Evolutionary Biology, and Earth and Planetary Sciences, Harvard University, 20 Oxford Street, Cambridge, Massachusetts 02138, USA*
‡ *Department of Statistics, University of Washington, Seattle, Washington 98195, USA*
§ *Anthropologisches Institut & Museum, Universität Zürich-Irchel, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*
‖ *Human Origins Group, The Natural History Museum, Cromwell Road, London SW7 5BD, UK*
¶ *Academic Affairs, Field Museum, Chicago, Illinois 60605-2496, USA*

Divergence times estimated from molecular data often considerably predate the earliest known fossil representatives of the groups studied. For the order Primates, molecular data calibrated with various external fossil dates uniformly suggest a mid-Cretaceous divergence from other placental mammals, some 90 million years (Myr) ago[1–9], whereas the oldest known fossil primates are from the basal Eocene epoch (54–55 Myr ago). The common ancestor of primates should be earlier than the oldest known fossils[10,11], but adequate quantification is needed to interpret possible discrepancies between molecular and palaeontological estimates. Here we present a new statistical method, based on an estimate of species preservation derived from a model of the diversification pattern, that suggests a Cretaceous last common ancestor of primates, approximately 81.5 Myr ago, close to the initial divergence time inferred from molecular data. It also suggests that no more than 7% of all primate species that have ever existed are known from fossils. The approach unites all the available palaeontological methods of timing evolutionary events: the fossil record, extant species and clade diversification models.

Although several molecular studies indicate that the lineage leading to primates diverged from other eutherian mammals about 90 Myr ago, diagnostic morphological features of primates possibly emerged later, potentially explaining why recognizable

Table 1 Relative sampling intensities for the primate fossil record

Epoch	k	T_k	Observed number of species, D_k	Relative sampling intensity, p_k	
				Scheme 1	Scheme 2
Late Pleistocene	1	0.15	19	1.0	1.0
Middle Pleistocene	2	0.9	28	1.0	1.0
Early Pleistocene	3	1.8	22	1.0	1.0
Late Pliocene	4	3.6	47	1.0	1.0
Early Pliocene	5	5.3	11	1.0	0.5
Late Miocene	6	11.2	38	1.0	0.5
Middle Miocene	7	16.4	46	1.0	1.0
Early Miocene	8	23.8	36	1.0	0.5
Late Oligocene	9	28.5	4	1.0	0.1
Early Oligocene	10	33.7	20	1.0	0.5
Late Eocene	11	37.0	32	1.0	1.0
Middle Eocene	12	49.0	103	1.0	1.0
Early Eocene	13	54.8	68	1.0	1.0
Pre-Eocene	14		0	0.1	0.1

Data are shown for a total of 235 modern species. References for the data can be found in the Supplementary Information.
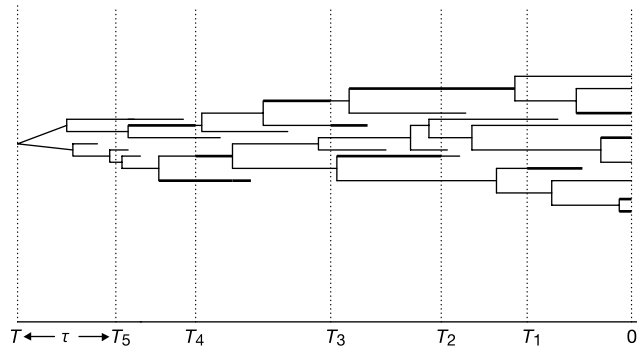
Figure 1 An illustration of the stochastic model of fossil finds. Bases of five stratigraphic intervals at $T_1, ..., T_5$ Myr ago are shown along the x axis. The temporal gap between the base of the final interval and the point at which the two founding species originate is denoted by τ. Thick lines indicate species found in the fossil record. Time 0 is the present day.

primates appear late in the known fossil record. However, divergence between strepsirrhines (lemurs and lorises) and haplorhines (tarsiers and anthropoids), postdating development of defining morphological features of primates, probably occurred soon after primates diverged from other mammals[5,10]. We estimate the divergence time of strepsirrhines and haplorhines—that is, the time of the last common ancestor (LCA) of living primates—using data from the fossil record.

Existing statistical methods designed to account for incompleteness of the fossil record typically use the size and distribution of gaps within observed stratigraphic ranges of lineages to estimate the size of the temporal gap between the oldest fossil and the LCA of the lineages[12]. Although these methods are useful for species already known from the fossil record, they are inappropriate for estimating the time of the LCA of higher taxonomic groups because they cannot account for species not preserved at all[10,12]. Our method for estimating the temporal gap uses an estimate of the proportion of species from the group actually preserved in the fossil record and the shape of its diversification curve. The speciation model we use is the non-homogeneous Markov branching process[13,14] (Fig. 1).

We have developed a computational approach for estimating the length of the temporal gap between the oldest known fossils and the LCA of a taxonomic group, as well as an estimated confidence interval (see Methods). We use as input the number of extant species, the mean species lifetime, the ages of the bases of the relevant stratigraphic intervals, the numbers of fossil species found in those intervals and the relative sizes of the sampling intensities in each interval. Estimates of the absolute values of the sampling intensities may also be found (see Methods, equation (6)). This gives an estimate of the proportion of species that existed in an interval that were found as fossils.

Different diversification models can be explored with our approach, but we present only results for a logistic diversification model. Logistic growth is the most biologically realistic model[15], matching the general expectation of an equilibrium diversity level. In our case, equilibrium diversity is achieved not through an increase in per-species extinction rate coupled with a decrease in per-species origination rate[16,17], but solely through a decrease in origination rate; our extinction rate is independent of standing diversity[18,19]. We parametrize logistic growth by the time at which diversity reached 90% of its present-day value. In the absence of any simple means of assessing relative sizes of sampling intensities for the stratigraphic intervals (although see refs 20 and 21), we fitted our model to two schemes (see Table 1 for details). In both schemes, the sampling intensity in the interval preceding the first to yield fossil primates was one-tenth of that in the first such interval. This option reflects the possibility that the first members of the clade were unusually small[22], perhaps had very low population sizes and/ or had a limited geographical distribution and may thus be harder to find in the fossil record. It has been demonstrated that preservation rates for mammals were lower in the Cretaceous period than during the Cenozoic era[20].

Various parametrizations of the logistic diversification model were investigated, but the great diversity of holarctic primates during the Eocene suggests that 90% of modern diversity had already been reached by the Middle Eocene, 49 Myr ago. All analyses were run with the mean species duration set to 2.5 Myr (ref. 23), but the results change by only 1% if mean longevity is increased to 3 Myr or decreased to 2 Myr, so the method appears to be relatively insensitive to assigned species longevity.

To address how well our model describes the data, we used the fitted values of the X^2 statistic defined in equation (5) (see Methods). Scheme 1 gave a value of $X^2 = 116.8$, considerably larger than expected. Because of this lack of fit, this scheme is not discussed further. Scheme 2 gave a value of $X^2 = 28.5$, consistent with expectations; this scheme provides an adequate fit to the data. The estimated time of the LCA for Scheme 2 is 81.5 Myr ago with a 95% confidence interval of (72.0, 89.6) Myr ago. Estimates of the time of the LCA of primates based on 90% of modern diversity being reached by the base of the Miocene epoch are also considerably older than the oldest known fossil, lying even further in the past. The sampling probabilities up to the base of the Eocene have an average value of 0.057, with an upper 95% confidence limit of 0.074.

While our results agree broadly with a molecular estimate of the time of the strepsirrhine and haplorhine divergence[5], they contradict widely accepted palaeontological estimates. Gingerich and Uhen[24] argued that, at a 95% confidence level, primates originated 55–63 Myr ago. However, their model gives the same results regardless of the number of modern primate species and regardless of species preservation rate. In fact, application of that model to a 6 Myr gap in the primate fossil record between the Early and the Late Oligocene epoch yields an extremely low probability for the existence of primates during that gap. Foote *et al.*[20] argue that, if molecular clock estimates of Cretaceous origins of living mammalian orders are correct, the preservation potential per lineage per million years (r) must be at least an order of magnitude smaller than they had estimated: between 0.25 and 0.37 (refs 20 and 25) for Cenozoic mammals and 0.03 for Cretaceous mammals. Our sampling intensities (α_j) divided by average species longevity are equivalent to r of ref. 20 and may be estimated from equation (6) (see Methods). The average values for preservation potential obtained using our approach—0.023 per lineage per Myr, and 0.003 per lineage per Myr for the last interval—are, indeed, an

order of magnitude smaller. Preservation rates proposed[20] for modern eutherian mammals are based either entirely (Cenozoic) or predominantly (Cretaceous) on North American faunas. North America is the best-sampled region in the world, and estimates based on that region will necessarily overestimate preservation rates of groups with a wider distribution.

The preservation rates of ref. 20 are likely to be overestimates for several reasons. Methods for assessing completeness based exclusively on the fossil record can only account for gaps within known lineages. Foote[25] demonstrated that the method used in ref. 20 will overestimate preservation potential where chronological gaps occur. The primate fossil record as a whole has several large gaps, most notably a general absence of fossils between the Early and Late Oligocene (at least 6 Myr) and the absence of a fossil record for the Malagasy lemurs, a diverse group with at least 46 living and subfossil species. Geographical gaps are equally substantial. Living primates are essentially confined to tropical and subtropical habitats[26]. Primates populated substantial parts of the northern continents only when these areas supported subtropical habitats, during the Eocene and the Miocene. Yet 47% of known fossil primate species come from North America and Europe and, for the first half of palaeontologically documented primate evolution, sites yielding fossil primates are largely restricted to these two regions (Fig. 2).
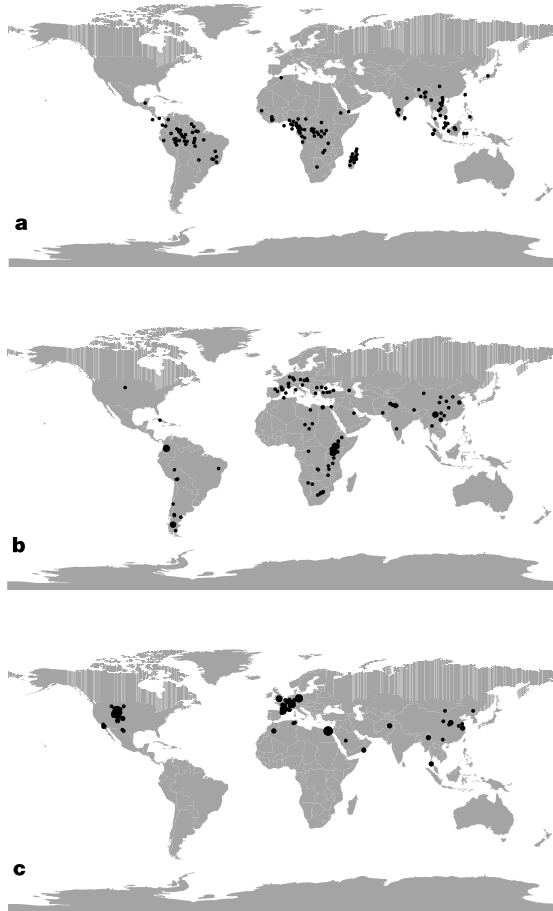

Figure 2 Mid-range of geographical distribution for individual modern and fossil primate species. **a**, Modern and sub-fossil primates (170 species obtained from Wolfheim's review of the distribution of modern primates[28]). **b**, Fossil species for the Late Pleistocene to the Late Oligocene (167 species). **c**, Fossil species from the Early Oligocene to the Early Eocene (196 species). (The database of fossil primates was compiled from a large number of published sources. A full list of references can be found at http://www.unizh.ch/anthro/Main/Who/Soligo/supinfo2.html).

Direct reading of the known fossil record suggests that primates originated during the Palaeocene in the northern continents and subsequently migrated southwards. An alternative interpretation is that primates originated earlier in the poorly documented southern continents and expanded northwards when climatic conditions permitted. □

Methods

A model for speciation

The speciation model we use is the non-homogeneous Markov branching process[13,14]. To model evolution after the LCA, we start with two species at time 0. Species have exponential lifetimes with mean $1/\lambda$, time being measured in millions of years. A species that goes extinct at time u is replaced by an average of $m(u)$ new species. Write Z_t for the number of species alive at time t, and define $B[s,t]$ to be the number of species that are born in the time interval $[s,t)$. The expected number of species extant at time t is given[13] by

$$EZ_t = 2 \exp\left\{ \lambda \int_0^t (m(u) - 1)\mathrm{d}u \right\} \tag{1}$$

and

$$EB[s, t] = \lambda \int_s^t m(u)EZ_u\mathrm{d}u, \quad s < t \tag{2}$$

To model the mean diversification, we use the logistic function for which $EZ_t =]2/\{\gamma + (1 - \gamma)e^{-\rho t}\}$.

A model for fossil finds

We divide time into k stratigraphic intervals, beginning from the present and proceeding into the past (see Table 1 and Fig. 1). The base of the first (youngest) stratigraphic interval is at T_1 Myr ago and the base of the kth at T_k Myr ago. The earliest fossil is found in this interval. The founding species originate at time $T = T_1 + \tau$ Myr ago, and we define a $(k + 1)$th stratigraphic interval that has its based at $T_{k+1} := T$ Myr ago and ends T_k Myr ago. No fossils have been found in this interval. We estimate the parameter τ, the temporal gap, using as data the number of different species found in the fossil record in the first, second, …, kth intervals.

To do this, we model the number of species alive u Myr ago by the value Z_{T-u} of the Markov branching process. The number N_j of distinct species living in the jth stratigraphic interval having base T_j Myr ago is the sum of those that were extant at the beginning of the interval, Z_{T-T_j}, plus those that originated in the interval $B[T - T_j, T - T_{j-1}]$. It follows from equations (1) and (2) that the expected number of distinct species that can be sampled in the jth stratigraphic interval is

$$EN_j = EZ_{T-T_{j-1}} + \lambda \int_{T-T_j}^{T-T_{j-1}} EZ_u\mathrm{d}u, \quad j = 1, \dots, k+1 \tag{3}$$

We assume that, conditional on the number of distinct species N_j in the jth stratigraphic interval ($j = 1, 2, \dots, k + 1$), the number of species D_j actually found in the fossil record in this interval is a binomial random variable with parameters N_j and α, $j = 1, 2, \dots, k$. Furthermore, the D_j are assumed to be conditionally independent given the N_j. The parameter α_j gives the probability of sampling a species in the jth stratigraphic interval. Under this sampling model, the expected number of species found in the jth interval is

$$ED_j = \alpha_j EN_j, \quad j = 1, \dots, k+1 \tag{4}$$

Statistical approach

Our method estimates τ by minimizing a weighted sum of squares of differences between the observed numbers of species found in the jth stratigraphic interval and the expected numbers given by equation (4). We assume that the number of species N_0 alive now is equal to the expected number under the diversification model; thus $N_0 = EZ_T$, which serves to determine one parameter of the diversification model.

Assuming that α_j is small, it can be shown that the variance of D_j is approximately $\alpha_j EN_j$. Our statistic therefore takes the form

$$X^2 = \sum_{j=1}^{k+1} (D_j - \alpha_j EN_j)^2 / \alpha_j EN_j \tag{5}$$

where $D_{k+1} = 0$ because no species are found from the earliest stratigraphic interval.

We also model the form of the sampling probabilities, α_j. One parsimonious choice is to set

$$\alpha_j = \alpha p_j, \quad j = 1, 2, \dots, k+1 \tag{6}$$

where the p_j are known relative sampling intensities and the scale parameter α is to be estimated. Our estimates of τ and α are given by the values that minimize the quantity X^2 defined in equation (5). Estimates of the absolute sampling fractions can be obtained from equation (6) using the estimate of α. We can also estimate the average species longevity, but for simplicity we assume it is known here.

Bias and approximate confidence intervals

We can obtain approximate confidence intervals for τ and α using a parametric bootstrap approach; compare with Ch. 5 of ref. 27. We simulate b realizations of a birth–death

process Z_t starting from $Z_0 = 2$, using the values τ_0 and α_0 of τ and α estimated from the data as the parameters in each run. Paths in which either branch dies out before time T are excluded from the analysis. For each accepted run, i, we re-estimate τ and α, getting values $\tau_i, \alpha_i, i = 1, \ldots, b$. Run i also produces a value X_i^2 of the statistic in equation (5). Assuming[27] that the values of $\tau_i - \tau_0$ mimic the distribution of $\tau_0 - \tau$, we obtain a bias-corrected estimate of τ of $2\tau_0 - \bar{\tau}$, where $\bar{\tau}$ is the sample average of the replicates $\tau_b \ldots, \tau_b$. An approximate $100(1 - 2v)\%$ equal-tailed confidence interval for τ is then given by $(2\tau_0 - \tau_{(b(1-v))}, 2\tau_0 - \tau_{(bv)})$, where $\tau_{(j)}$ is the jth largest of $\tau_b \ldots, \tau_b$. We used $b = 2,500$. A similar method is used to find an upper 95% confidence interval for α and the α_f. The empirical distribution of X^2 can be found from the values of $X_b^2, \ldots, X_b^2$.

1. Janke, A., Feldmaier-Fuchs, G., Thomas, W. K., von Haeseler, A. & Pääbo, S. The marsupial mitochondrial genome and the evolution of placental mammals. *Genetics* **137**, 243–256 (1994).
2. Easteal, S., Collett, C. & Betty, D. *The Mammalian Molecular Clock* (Landes, Austin, 1995).
3. Hedges, S. B., Parker, P. H., Sibley, C. G. & Kumar, S. Continental breakup and the ordinal diversification of birds and mammals. *Nature* **381**, 226–229 (1996).
4. Springer, M. S. *et al.* Endemic African mammals shake the phylogenetic tree. *Nature* **388**, 61–64 (1997).
5. Arnason, U., Gullberg, A. & Janke, A. Molecular timing of primate divergences as estimated by two nonprimate calibration points. *J. Mol. Evol.* **47**, 718–727 (1998).
6. Huelsenbeck, J. P., Larget, B. & Swofford, D. A compound Poisson process for relaxing the molecular clock. *Genetics* **154**, 1879–1892 (2000).
7. Madsen, O. *et al.* Parallel adaptive radiations in two major clades of placental mammals. *Nature* **409**, 610–614 (2001).
8. Murphy, W. J. *et al.* Molecular phylogenetics and the origins of placental mammals. *Nature* **409**, 614–618 (2001).
9. Eizirik, E., Murphy, W. J. & O'Brien, S. J. Molecular dating and biogeography of the early placental mammal radiation. *J. Hered.* **92**, 212–219 (2001).
10. Martin, R. D. Primate origins: plugging the gaps. *Nature* **363**, 223–234 (1993).
11. Martin, R. D. in *Major Topics in Primate and Human Evolution* (eds Wood, B. A., Martin, L. B. & Andrews, P.) 1–31 (Cambridge Univ. Press, Cambridge, 1986).
12. Marshall, C. R. in *The Adequacy of the Fossil Record* (eds Donovan, S. K. & Paul, C. R. C.) 23–53 (Wiley, London, 1998).
13. Harris, T. E. *The Theory of Branching Processes* (Springer, Berlin, 1963).
14. Nee, S., May, R. M. & Harvey, P. H. The reconstructed evolutionary process. *Phil. Trans. R. Soc. Lond. B* **344**, 305–311 (1994).
15. Raup, D. M., Gould, S. J., Schopf, T. M. & Simberloff, D. S. Stochastic models of phylogeny and the evolution of diversity. *J. Geol.* **81**, 525–542 (1973).
16. MacArthur, R. H. & Wilson, E. O. An equilibrium theory of insular zoogeography. *Evolution* **17**, 373–387 (1963).
17. MacArthur, R. H. & Wilson, E. O. *The Theory of Island Biogeography* (Princeton Univ. Press, Princeton, New Jersey, 1967).
18. Sepkoski, J. J. Jr A kinetic model of Phanerozoic taxonomic diversity: I. Analysis of marine orders. *Paleobiology* **4**, 223–225 (1978).
19. Alroy, J. in *Biodiversity Dynamics* (eds McKinney, M. L. & Drake, J. A.) 232–287 (Columbia Univ. Press, New York, 1998).
20. Foote, M., Hunter, J. P., Janis, C. M. & Sepkoski, J. J. Jr Evolutionary and preservational constraints on origins of biological groups: divergence times of eutherian mammals. *Science* **283**, 1310–1314 (1999).
21. Foote, M. & Sepkoski, J. J. Absolute measures of the completeness of the fossil record. *Nature* **398**, 415–417 (1999).
22. Alroy, J. Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science* **280**, 731–733 (1998).
23. Alroy, J. Quantitative mammalian biochronology and biogeography of North America. PhD dissertation. Univ. Chicago (1994).
24. Gingerich, P. D. & Uhen, M. D. Time of origin of primates. *J. Hum. Evol.* **27**, 443–445 (1994).
25. Foote, M. Estimating taxonomic durations and preservation probability. *Paleobiology* **23**, 278–300 (1997).
26. Martin, R. D. *Primate Origins and Evolution: A Phylogenetic Reconstruction* (Chapman Hall/Princeton Univ. Press, London/New Jersey, 1990).
27. Davison, A. C. & Hinkley, D. V. *Bootstrap Methods and their Applications* (Cambridge Univ. Press, Cambridge, 1997).
28. Wolfheim, J. H. *Primates of the World: Distribution Abundance and Conservation* (Univ. Washington Press, Seattle/London, 1983).

Competing interests statement

The authors declare that they have no competing financial interests.

Fisheries productivity in the northeastern Pacific Ocean over the past 2,200 years

Bruce P. Finney*†, **Irene Gregory-Eaves**†‡, **Marianne S. V. Douglas**§ & **John P. Smol**‡

* *Institute of Marine Science, University of Alaska Fairbanks, Fairbanks, Alaska 99775-7220, USA*
‡ *Paleoecological Environmental Assessment and Research Laboratory (PEARL), Department of Biology, Queen's University, Kingston, Ontario K7L 3N6, Canada*
§ *Paleoenvironmental Assessment Laboratory (PAL), Department of Geology, University of Toronto, 22 Russell St., Toronto, Ontario M5S 3B1, Canada*
† *These authors contributed equally to this work*

Historical catch records suggest that climatic variability has had basin-wide effects on the northern Pacific and its fish populations, such as salmon, sardines and anchovies[1–7]. However, these records are too short to define the nature and frequency of patterns. We reconstructed ~2,200-year records of sockeye salmon abundance from sediment cores obtained from salmon nursery lakes on Kodiak island, Alaska. Large shifts in abundance, which far exceed the decadal-scale variability recorded during the past 300 years[1–8], occurred over the past two millennia. A marked, multi-centennial decline in Alaskan sockeye salmon was apparent from ~100 BC to AD 800, but salmon were consistently more abundant from AD 1200 to 1900. Over the past two millennia, the abundances of Pacific sardine and Northern anchovy off the California coast, and of Alaskan salmon, show several synchronous patterns of variability. But sardines and anchovies vary out of phase with Alaskan salmon over low frequency, which differs from the pattern detected in historical records[5,6]. The coherent patterns observed across large regions demonstrate the strong role of climatic forcing in regulating northeastern Pacific fish stocks.

Salmon are important ecological, economical and cultural resources in the northern Pacific region, and their response to future climatic change is very uncertain[9]. Long-term relationships between sockeye salmon populations and climatic change can be evaluated by analysing sediment cores from their nursery lakes[8]. After one to four years of feeding in the northern Pacific, sockeye salmon (*Oncorhyncus nerka*) return to their natal lake/stream system to spawn and die[10]. The nutrients derived from spawned carcasses can be significant relative to other sources, and may be reconstructed from palaeolimnological records of $\delta^{15}N$ and algal bioindicators, such as diatoms[8]. Periods of greater input of salmon-derived nutrients (SDN), and hence greater sockeye salmon abundance, corresponded to higher sedimentary $\delta^{15}N$ and more eutrophic diatom taxa. Here we show reconstructions of salmon abundance over the past two millennia from lakes on Kodiak island, Alaska (Fig. 1), where our proxies have been successfully calibrated with monitoring data[8]. This region is one of the most important salmon-producing areas of the northern Pacific, and historical records suggest that its salmon abundances are representative of population trends in Alaska[4].

In Karluk lake (57° 25′ N, 154° 05′ W), our sedimentary indicators of SDN—$\delta^{15}N$ and diatoms—show striking changes and a strong degree of coherence ($r^2 = 0.79$, $n = 100$, $P < 0.001$) over the past ~2,200 years (Fig. 2). In the oldest sediments, around 200 BC, we infer from the high $\delta^{15}N$ values and the strong presence of mesotrophic to eutrophic diatom taxa (for example, *Stephanodiscus minutulus/parvus*) that the return of sockeye salmon to Karluk lake was high, and similar to levels (~3 million yr^{-1}) present when