Bahama Bank, in *Proceedings of the Ocean Drilling Program, Scientific Results*, P. K. Swart, G. P. Eberli, M. J. Malone, J. F. Sarg, Eds. (Ocean Drilling Program, College Station, TX, 2000), vol. 166, pp. 137–143.
13. J. Berry, O. Bjorkman, *Annu. Rev. Plant Physiol.* **31**, 491 (1980).
14. R. J. Ellis, *Nature* **463**, 164 (2010).
15. G. N. Somero, *Annu. Rev. Physiol.* **57**, 43 (1995).
16. H. O. Pörtner, *Comp. Biochem. Physiol.* **132**, 739 (2002).
17. H. O. Pörtner, *Naturwissenschaften* **88**, 137 (2001).
18. P. B. Wignall, R. J. Twitchett, *Spec. Pap. Geol. Soc. Am.* **356**, 395 (2002).
19. T. Galfetti *et al.*, *Sediment. Geol.* **204**, 36 (2008).
20. H.-O. Pörtner, *J. Exp. Biol.* **213**, 881 (2010).
21. J. M. Callaway, D. B. Brinkman, *Can. J. Earth Sci.* **26**, 1491 (1989).
22. C. B. Cox, D. G. Smith, *Geol. Mag.* **110**, 405 (1973).
23. P. B. Wignall, R. Morante, R. Newton, *Geol. Mag.* **135**, 47 (1998).
24. A. H. Knoll, R. K. Bambach, J. L. Payne, S. Pruss, W. W. Fischer, *Earth Planet. Sci. Lett.* **256**, 295 (2007).
25. S. G. Lucas, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **143**, 347 (1998).
26. M. Borsuk-Białynicka, E. Cook, S. E. Evans, T. Maryań, *Acta Palaeontol. Pol.* **44**, 167 (1999).
27. H.-D. Sues, N. C. Fraser, *Triassic Life on Land* (Columbia Univ. Press, New York, 2010).
28. T. Galfetti *et al.*, *Geology* **35**, 291 (2007).
29. E. Schneebeli-Hermann *et al.*, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **339–341**, 12 (2012).
30. C. V. Looy, W. A. Brugman, D. L. Dilcher, H. Visscher, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13857 (1999).
31. H. O. Pörtner, R. Knust, *Science* **315**, 95 (2007).
32. M. J. Angilletta, *Thermal Adaptation- A Theoretical and Empirical Synthesis* (Oxford Univ. Press, New York, 2009).
33. J. A. Sheridan, D. Bickford, *Nat. Clim. Change* **1**, 401 (2011).
34. B. Metcalfe, R. J. Twitchett, N. Price-Lloyd, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **308**, 171 (2011).
35. R. J. Twitchett, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **154**, 27 (1999).
36. G. Chapelle, L. S. Peck, *Nature* **399**, 114 (1999).
37. S. M. Stanley, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 15264 (2009).
38. M. J. Orchard, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **252**, 93 (2007).
39. J. Chen, in *Mass Extinction and Recovery: Evidences from the Palaeozoic and Triassic of South China*, J. Rong, Z. Fang, Eds. (Univ. of Science and Technology of China Press, Heifei, 2004), vol. II, pp. 647–700.
40. H. Svensen *et al.*, *Earth Planet. Sci. Lett.* **277**, 490 (2009).
41. S. M. Stanley, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 19185 (2010).
42. G. J. Retallack, E. S. Krull, *Aust. J. Earth Sci.* **46**, 785 (1999).
43. W. Broecker, S. Peacock, *Global Biogeochem. Cycles* **13**, 1167 (1999).
44. M. Huber, *Science* **321**, 353 (2008).
45. A. M. Ziegler, M. L. Hulver, D. B. Rowley, in *Late Glacial and Postglacial Environmental Changes–Quaternary, Carboniferous-Permian and Proterozoic*, I. P. Martini, Ed. (Oxford Univ. Press, New York, 1997), pp. 111–146.
46. G. Muttoni *et al.*, *Geoarabia* **14**, 17 (2009).
47. D. J. Lehrmann *et al.*, *Palaios* **18**, 138 (2003).
48. R. A. Locarnini *et al.*, *World Ocean Atlas 2009, Volume 1: Temperature*, S. Levitus, Ed. [National Oceanic and Atmospheric Administration (NOAA) Atlas NESDros. Inf. Serv. 68, U.S. Government Printing Office, Washington, DC, 2010].
49. H. Song *et al.*, *Geology* **39**, 739 (2011).
50. D. Sun, S. Shen, in *Mass Extinction and Recovery: Evidences from the Palaeozoic and Triassic of South China*, J. Rong, Z. Fang, Eds. (Univ. of Science and Technology of China Press, Heifei, China, 2004), vol. II, pp. 543–570.
51. H. Pan, D. H. Erwin, *Palaeoworld* **4**, 249 (1994).
52. L. O'Dogherty *et al.*, *Geodiversitas* **31**, 213 (2009).

Acknowledgments: D. Lutz, F. Nenning, B. Yang, and X. Liu are acknowledged for lab and field assistance. This study was supported by Chinese 973 Program (2011CB808800) and the Natural Science Foundation of China (41172024 and 40830212). Y.S. acknowledges China University of Geosciences and China Scholarship Council for split-site Ph.D. at Wuhan, Leeds, and Erlangen.

Supplementary Materials
www.sciencemag.org/cgi/content/full/338/6105/366/DC1
Materials and Methods
Supplementary Text
Fig. S1
Tables S1 to S4
References (*53–150*)

1 May 2012; accepted 4 September 2012
10.1126/science.1224126

A Complete Terrestrial Radiocarbon Record for 11.2 to 52.8 kyr B.P.

Christopher Bronk Ramsey,[1]* Richard A. Staff,[1] Charlotte L. Bryant,[2] Fiona Brock,[1] Hiroyuki Kitagawa,[3] Johannes van der Plicht,[4,5] Gordon Schlolaut,[6] Michael H. Marshall,[7] Achim Brauer,[6] Henry F. Lamb,[7] Rebecca L. Payne,[8] Pavel E. Tarasov,[9] Tsuyoshi Haraguchi,[10] Katsuya Gotanda,[11] Hitoshi Yonenobu,[12] Yusuke Yokoyama,[13] Ryuji Tada,[13] Takeshi Nakagawa[8]

Radiocarbon (^{14}C) provides a way to date material that contains carbon with an age up to ~50,000 years and is also an important tracer of the global carbon cycle. However, the lack of a comprehensive record reflecting atmospheric ^{14}C prior to 12.5 thousand years before the present (kyr B.P.) has limited the application of radiocarbon dating of samples from the Last Glacial period. Here, we report ^{14}C results from Lake Suigetsu, Japan (35°35′N, 135°53′E), which provide a comprehensive record of terrestrial radiocarbon to the present limit of the ^{14}C method. The time scale we present in this work allows direct comparison of Lake Suigetsu paleoclimatic data with other terrestrial climatic records and gives information on the connection between global atmospheric and regional marine radiocarbon levels.

Lake Suigetsu contains annually laminated sediments that preserve both paleoclimate proxies and terrestrial plant macrofossils that are suitable for radiocarbon dating. The lake's potential to provide an important archive of atmospheric radiocarbon (^{14}C) was realized in 1993 (*1*). However, the single SG93 sediment core then recovered included missing intervals between successive sections (*2*). This, together with the difficulty of visual varve counting, resulted in inconsistency between the SG93 and other ^{14}C calibration records (*3*). The SG06 core-set recovered in 2006 consists of four parallel cores that together avoid any such sedimentary gaps (*4*). Here, we report 651 ^{14}C measurements covering the period between 11.2 and 52.8 thousand years before the present (kyr B.P.) tied to a time scale derived from varve counting and temporal constraints from other records. Using visual markers, we applied a composite depth (CD) scale to all cores, including SG93. We also define an event-free depth (EFD), which is the CD with substantial macroscopic event layers (such as turbidites and tephras) removed.

Accelerator mass spectrometry radiocarbon dating (*5*) has been conducted on terrestrial plant macrofossils selected from the SG06 cores to cover the full ^{14}C time range, from the present to the detection limit of the ^{14}C method (0 to 41 m CD) (table S1). The results already reported from the control period (0 to 12.2 kyr B.P.) (*6*), covered by the tree-ring–derived calibration curve (*7*), act to demonstrate the integrity of the sediments and to anchor the floating SG06 varve chronology, because varves do not extend into the Holocene.

The varve-based chronology for SG06 (*5*, *8*, *9*) provides our best estimate of the true age of the cores for the period ~10.2 to 40.0 kyr B.P., based only on information from the site. It provides good relative chronological precision and has the advantage of being independent of other dating techniques. However, the cumulative counting uncertainty inevitably increases with age (~6% at 40 kyr B.P.). The full varve chronology (Fig. 1A and table S1) has been extrapolated on the basis of EFD to cover the period 40 to 53 kyr B.P.

To better constrain the uncertainties in the varve chronology, we can directly compare the Suigetsu data set and other archives that provide information on atmospheric ^{14}C and associated independent ages. The two most useful records for this purpose are the Bahamas speleothem GB89-25-3 (*10*) and the Hulu Cave speleothem H82 (*11*), both of which have extensive ^{14}C- and U-Th–based chronologies. In both cases, we would expect the radiocarbon in the speleothems to respond to changes in atmospheric ^{14}C content, despite the groundwater containing a dead-carbon fraction (DCF) from dissolved carbonates. Estimated DCF for these speleothems was 2075 ± 270 radiocarbon

[1]University of Oxford, Oxford, UK. [2]Natural Environment Research Council Radiocarbon Facility, Scottish Universities Environmental Research Centre, East Kilbride, UK. [3]Nagoya University, Nagoya, Japan. [4]University of Groningen, Groningen, Netherlands. [5]University of Leiden, Leiden, Netherlands. [6]GeoForschungsZentrum German Research Centre for Geosciences, Potsdam, Germany. [7]Aberystwyth University, Aberystwyth, UK. [8]University of Newcastle, Newcastle upon Tyne, UK. [9]Free University Berlin, Berlin, Germany. [10]Osaka City University, Osaka, Japan. [11]Chiba University of Commerce, Chiba, Japan. [12]Naruto University of Education, Naruto, Japan. [13]University of Tokyo, Tokyo, Japan.

*To whom correspondence should be addressed. E-mail: christopher.ramsey@rlaha.ox.ac.uk

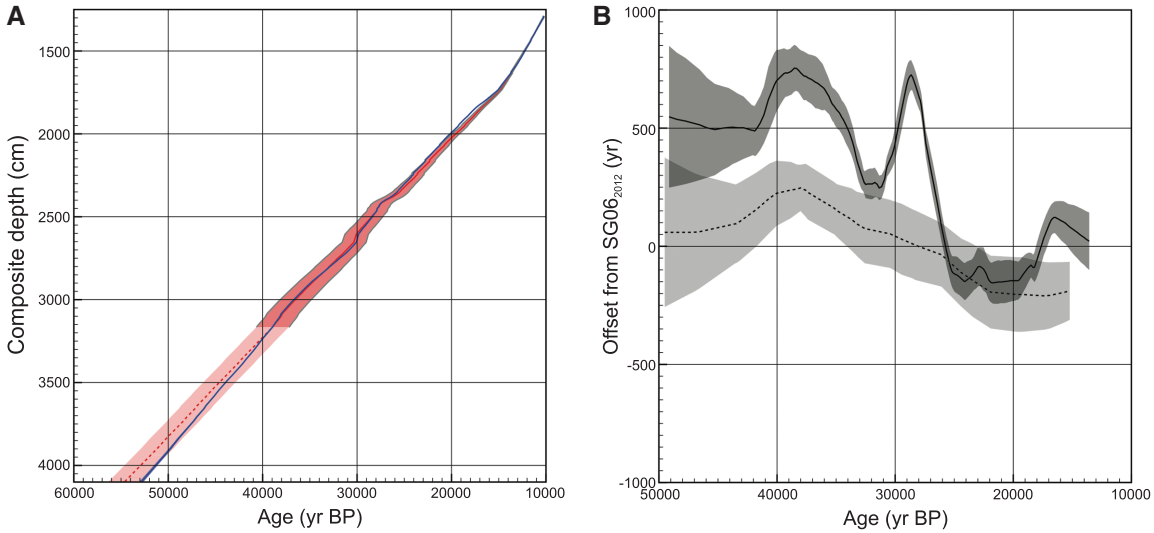
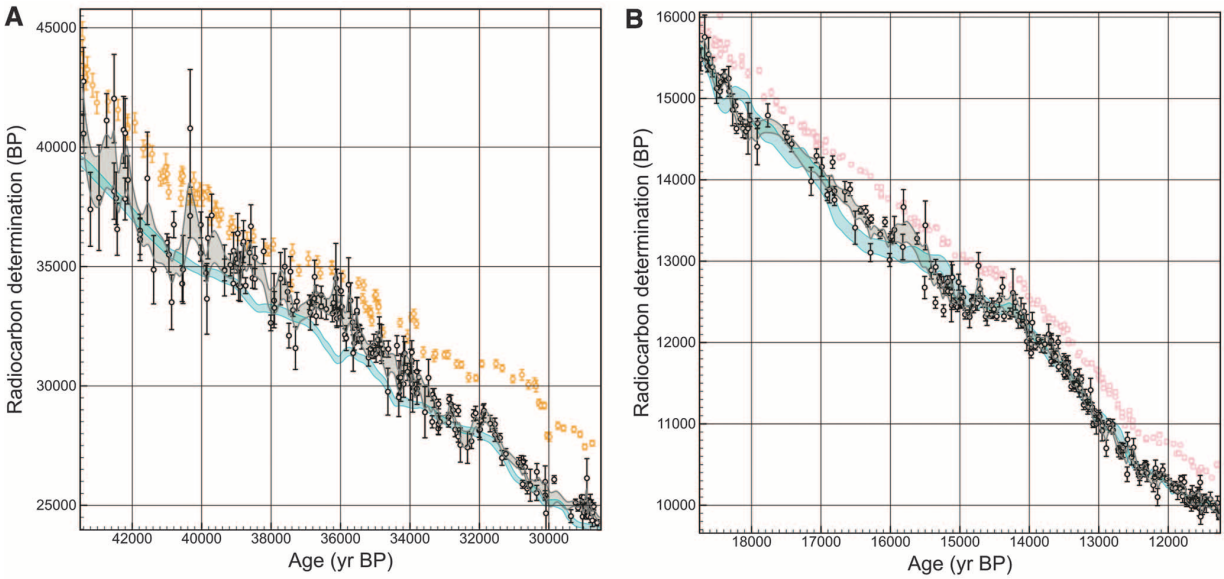


Fig. 1. (**A**) Age-depth model for SG06, based on varves (red), extrapolated by EFD (pink), and constrained by speleothems H82 and GB89-25-3 (blue). (**B**) Inferred differences in the age models for the Cariaco Basin (*22*) (solid curves) and Iberian Margin (*23*) (dotted curves) data sets in IntCal09 (*7*) compared to the SG06$_{2012}$ modeled chronology (higher offset implies that the time scale is older); see figs. S9 and S10.



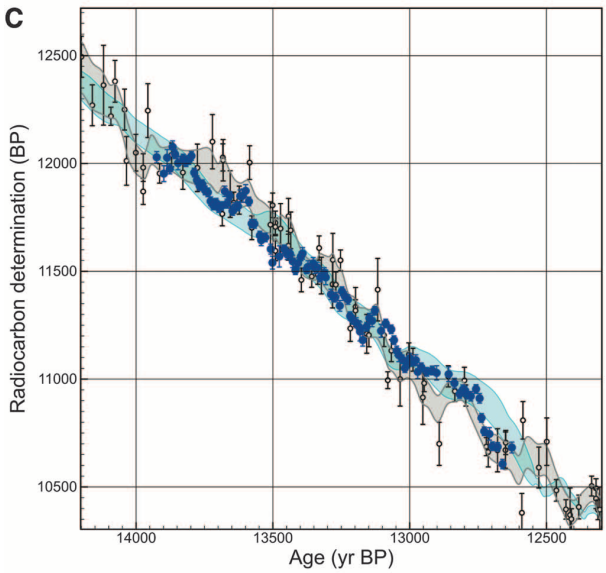


Fig. 2. Comparison of raw (black) and five-point moving mean (gray) Lake Suigetsu ^{14}C data against IntCal09 (*7*) (light blue), as well as (**A**) GB89-25-3 data (*10*) (orange), (**B**) H82 speleothem data (*11*) (pink) for the latter part of the H82 time scale, and (**C**) the LGP record (*17*) (blue). Whereas the match to the LGP data are good for most individual data points, the match would be improved if the SG06$_{2012}$ chronology were stretched by ~5% over this period (possible within errors). See also fig. S6. Error bars denote 1σ.

years (^{14}C yr) for GB89-25-3 (*10*) and 450 ± 70 ^{14}C yr for H82 (*11*), each at 1σ. We have modeled these two records onto the $SG06_{2012}$ varve chronology using a Poisson process model (*12*), which allows for nonlinear random deviation between the $SG06_{2012}$ varve chronology and the U-Th time scale underlying the speleothem records (*5*). The model provides independent estimates for the mean DCF of the speleothems of 2500 ± 90 for GB89-25-3 and 440 ± 25 for H82, in agreement with the initial estimates.

The model results can also be used to refine the SG06 chronology by including the constraints provided by the speleothem U-Th dates to 44 kyr B.P. This greatly reduces the uncertainty in the absolute chronology of the age-depth profile (Fig. 1A and table S1) and ensures that the SG06 data are on a U-Th–moderated time scale. There are some significant differences between the varve chronology and the model (notably in the period around 12.6 kyr B.P.), possibly associated with changes in the sedimentation rate and sedimentary processes at Suigetsu through that period; however, for most of the core, the agreement between the chronologies is good (Fig. 1A), with the ratio between the varve-only and model-inferred deposition rates being 1.01 ± 0.10.

The modeled chronology ($SG06_{2012}$ yr B.P.) is based on all of the available information from both Suigetsu and the other long records of atmospheric radiocarbon. This precise time scale (subcentennial 1σ uncertainties to ~25 kyr B.P.) allows us to place the detailed paleoclimate data from Suigetsu in a global context, facilitating the identification of potential leads and lags in climate change recorded in key paleoenvironmental archives.

The 651 terrestrial radiocarbon dates from the period 11,241 to 52,820 $SG06_{2012}$ yr B.P. provide a quasi-continuous record of atmospheric ^{14}C for this period (Fig. 2), for which the current calibration curve (*7*) is largely based on reservoir-affected marine data. As such, the Suigetsu calibration data set provides the first "backbone" upon which a comprehensive atmospheric calibration curve can be built. It will allow, for example, the archeology related to the extinction of Neanderthals and spread of anatomically modern humans into Europe to be calibrated against terrestrial reference data. With its high density of ^{14}C measurements, this data set also allows direct linking between SG06 and any other paleoenvironmental record with terrestrial radiocarbon data.

In addition to its importance for radiocarbon calibration and correlation between different climate records using radiocarbon, a full record of atmospheric ^{14}C has important implications for our understanding of the carbon cycle. The links between atmospheric ^{14}C and primary production are key, as is the connection between the atmosphere and the ocean. Figure 3 shows the inferred level of radiocarbon in the atmosphere (Δ^{14}C) at Lake Suigetsu, compared to the ^{10}Be record from Greenland (*13*, *14*). The peak in ^{10}Be production around 41 kyr B.P., assumed to be enhanced by the Laschamp geomagnetic excursion (*15*), is clearly visible in the Suigetsu data; the similar timing indicates congruence between the $SG06_{2012}$ and North Greenland Ice Core Project (NGRIP) GICC05 (*16*) time scales at this point.

Agreement between the Suigetsu radiocarbon record and those of the speleothems (GB89-25-3 and H82) is generally good (Fig. 2, A and B), though for the period 28-32 $SG06_{2012}$ kyr B.P., the implied reservoir offset for GB89-25-3 seems higher than for the older sections (Fig. 2A). Another ^{14}C record of great importance is the European Late Glacial Pine (LGP) record (*17*). Figure 2C shows the match of this record to that of Suigetsu. The match puts the younger end of the LGP sequence at 12627 ± 35 yr B.P. and gives the average radiocarbon offset of Suigetsu to the LGP data set as 8 ± 6 ^{14}C yr (older). Our modeling supports the fit provided by the linkage of the LGP to Tasmanian Huon Pine (*18*), rather than the suggested ^{10}Be-based fit to NGRIP (*19*).

We have used the radiocarbon sequence from Lake Soppensee, Switzerland, as an example of direct comparison with European varved lake sediments (*20*). This enables us to place the climate proxy signal associated with the Younger Dryas onset at Soppensee at 12607 ± 85 $SG06_{2012}$ yr B.P., some 156 ± 88 years later than cooling in Suigetsu, which, assuming congruence with GICCO5, probably lags the rapid cooling in Greenland (but only by 83 years, which could be synchronous within error margins). The match to Soppensee also enables us to place the European tephra marker of the Laacher See at 12842 ± 53 $SG06_{2012}$ yr B.P., which agrees well with the independent age estimate of 12880 ± 40 varve years (vyr) B.P. in Meerfelder Maar, Germany (*21*).

Direct comparison of the terrestrial radiocarbon signal from Suigetsu can also be made to that recorded in marine archives. The long records from the Cariaco Basin, west Atlantic (*22*), and the Iberian Margin record, northeast Atlantic (*23*), are climatically tuned to the Hulu Cave U-Th chronology. By constraining the total marine reservoir age for these two locations within reasonable limits (*5*), we are able to quantify offsets between the tuned chronologies and that of SG06

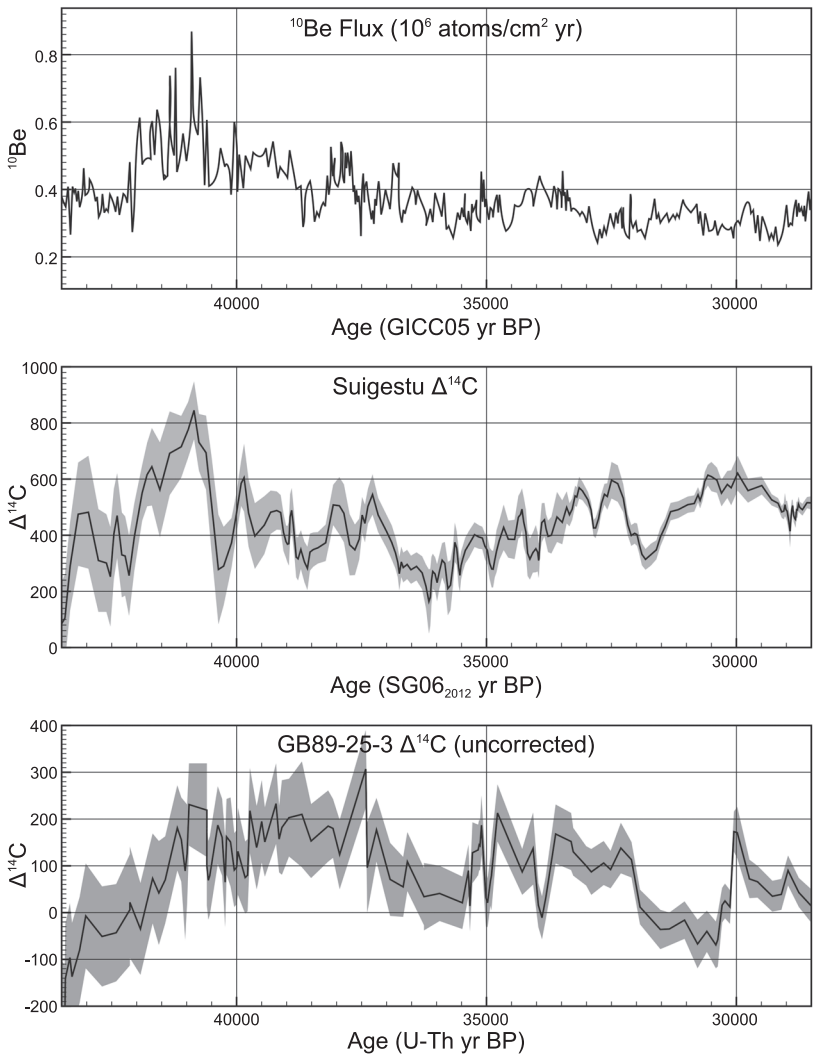


Fig. 3. Inferred Δ^{14}C values from Lake Suigetsu compared to data from GB89-25-3 (*10*) uncorrected for DCF. The ^{10}Be flux in Greenland (*13*, *14*) is shown for comparison.

(Fig. 1B). The Iberian Margin time scale is mostly consistent with SG06 (within uncertainties), except that it is ~200 years older at ~40 kyr B.P. and ~200 years younger at ~15 kyr B.P. The latter could be due to greater reservoir ages in the northeast Atlantic at the Last Glacial Maximum (LGM) than have previously been allowed. The Cariaco Basin time scale seems consistent with that for SG06 back to ~26 kyr B.P., at which point there is a discrete peak (amplitude ~500 years), perhaps due to specific choices in the climate-tuning. At ~40 kyr B.P., the Cariaco time scale, like that of the Iberian Margin, is older than that of $SG06_{2012}$, but by a much greater extent (~700 years). Given that both records are tuned to the same (Hulu) chronology, these differences are more likely to be due to lags in different regional environmental responses to global climate rather than offsets between the fundamental time scales of Hulu and GB89-25-3 (and, hence, $SG06_{2012}$) over this period.

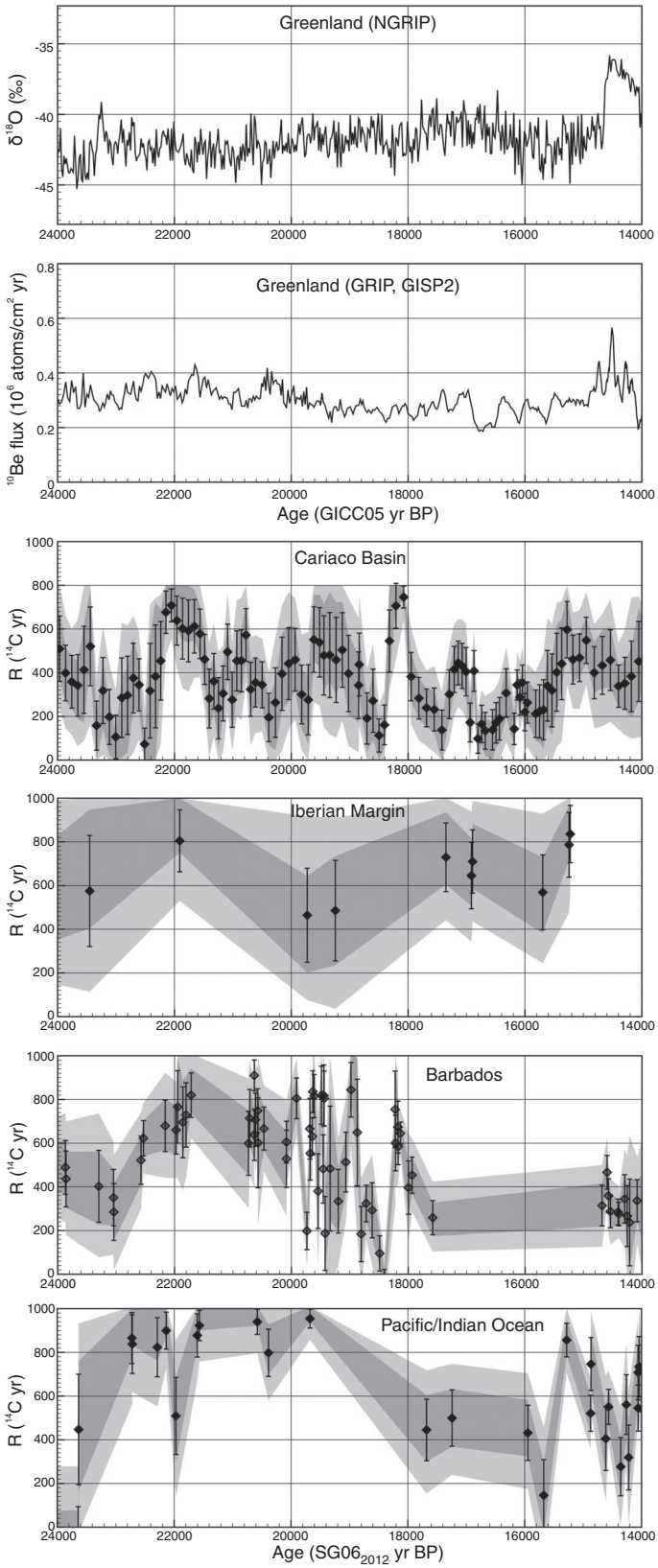


Fig. 4. Marine reservoirs relative to atmosphere (*R*), as deduced from the Suigetsu data on the modeled $SG06_{2012}$ chronology. For comparison, we show the ^{10}Be flux (*13*, *14*) and the δ^{18}O signal from Greenland on the GICC05 time scale (*16*).

We can also infer total marine reservoir age (*R*) from these models (Fig. 4 and fig. S11). In general, most of the signal we see for the Cariaco Basin is due to greater fluctuations in the Δ^{14}C of the atmosphere from that recorded in marine sediments. Some of the features mirror those of the ^{10}Be production rate in Greenland and might be related to ^{14}C production, others may reflect changes in local ocean dynamics. Although there are fewer data from the Iberian Margin, the rise in *R* between 16 and 15 kyr B.P. is still apparent, and there is a higher reservoir age inferred at ~22 kyr B.P., during the LGM.

For the marine coral data from IntCal09 (*7*), we additionally have direct U-Th dates; thus, *R* can be calculated more directly. Coral data, however, are not available at regular intervals, so we do not have a continuous record. Pacific coral data (*7*) in the range 30 to 39 kyr B.P. are consistently higher in Δ^{14}C than the terrestrial data from Suigetsu (fig. S12). We have to deduce that either the $SG06_{2012}$ time scale is substantially too young (a hypothesis not supported by the Iberian Margin data, identification of the Laschamp event as compared to Greenland, or the speleothem data used to correct the modeled $SG06_{2012}$ time scale) or some of the coral data have elevated U-Th ages or overestimated ^{14}C measurements (more likely, given the range of ^{14}C measurements on some coeval samples). It is clear that *R* for the Pacific and Barbados corals (*7*), as for the Cariaco Basin and the Iberian Margin foraminifera, is greater at ~22 kyr B.P. (Fig. 4)—possibly a consequence of lower global CO_2 during the LGM. In the period around 16 to 14 kyr B.P., there is a similar pattern in the Pacific/Indian Ocean as in Cariaco: a rapid rise in *R*, followed by a gradual fall.

The terrestrial radiocarbon record from Lake Suigetsu presented here, together with Holocene measurements (*6*), comprises 808 radiocarbon determinations from two core sites in the center of the lake, measured by three laboratories. Together, these give us a single, quasi-continuous record of purely atmospheric ^{14}C covering the full range of the radiocarbon technique. This will greatly benefit calibration of terrestrial radiocarbon samples in the period 12.5 to 52.8 kyr B.P. and will enable direct correlation between other key climate records and the Lake Suigetsu record itself, without any assumptions of climatic synchrony. An atmospheric record of ^{14}C over this whole time scale also facilitates a more comprehensive understanding of the long marine records in their oceanic context, rather than simply assuming that they represent atmospheric ^{14}C.

References and Notes

1. H. Kitagawa, J. van der Plicht, *Science* **279**, 1187 (1998).
2. R. A. Staff, C. Bronk Ramsey, T. Nakagawa, *Nucl. Instrum. Methods* **B268**, 960 (2010).
3. J. van der Plicht *et al.*, *Radiocarbon* **46**, 1225 (2004).
4. T. Nakagawa *et al.*, *Quat. Sci. Rev.* **36**, 164 (2012).
5. Materials and methods are available as supplementary materials on *Science* Online.
6. R. A. Staff *et al.*, *Radiocarbon* **53**, 511 (2011).
7. P. J. Reimer *et al.*, *Radiocarbon* **51**, 1111 (2009).
8. M. Marshall *et al.*, *Quat. Geochronol.* (2012).
9. G. Schlolaut *et al.*, *Quat. Geochronol.* (2012).
10. D. L. Hoffmann *et al.*, *Earth Planet. Sci. Lett.* **289**, 1 (2010).
11. J. Southon, A. L. Noronha, H. Cheng, L. R. Edwards, Y. Wang, *Quat. Sci. Rev.* **33**, 32 (2012).
12. C. Bronk Ramsey, *Quat. Sci. Rev.* **27**, 42 (2008).
13. R. Muscheler, personal communication.
14. GRIP and GISP2 ^{10}Be flux (*15*) are corrected onto the GICC05 time scale (*16*).
15. R. Muscheler, J. Beer, P. W. Kubik, H.-A. Synal, *Quat. Sci. Rev.* **24**, 1849 (2005).
16. A. Svensson *et al.*, *Clim. Past* **4**, 47 (2008).
17. B. Kromer *et al.*, *Radiocarbon* **46**, 1203 (2004).
18. Q. Hua *et al.*, *Quat. Sci. Rev.* **28**, 2982 (2009).
19. R. Muscheler *et al.*, *Nat. Geosci.* **1**, 263 (2008).
20. I. Hajdas *et al.*, *Clim. Dyn.* **9**, 107 (1993).
21. A. Brauer, C. Endres, J. F. W. Negendank, *Quat. Int.* **61**, 17 (1999).
22. K. Hughen, J. Southon, S. Lehman, C. Bertrand, J. Turnbull, *Quat. Sci. Rev.* **25**, 3216 (2006).
23. E. Bard, F. Rostek, G. Ménot-Combes, *Quat. Res.* **61**, 204 (2004).

Acknowledgments: This project was funded by the UK Natural Environment Research Council (NE/D000289/1 and NE/F003048/1, as well as Radiocarbon Facility support); the Deutsche Forschungsgemeinschaft (BR 2208/7-1 and TA 540/3-1); Ministry of Education, Culture, Sports, Science and Technology–Japan KAKENHI (21101002); and the John Fell Oxford University Press Research Fund.

Supplementary Materials

www.sciencemag.org/cgi/content/full/338/6105/370/DC1
Materials and Methods
Figs. S1 to S12
Tables S1 to S3
References (*24–29*)

28 June 2012; accepted 22 August 2012
10.1126/science.1226660

Genomic Variation in Seven Khoe-San Groups Reveals Adaptation and Complex African History

Carina M. Schlebusch,[1]*† Pontus Skoglund,[1]† Per Sjödin,[1] Lucie M. Gattepaille,[1] Dena Hernandez,[2] Flora Jay,[3] Sen Li,[1] Michael De Jongh,[4] Andrew Singleton,[2] Michael G. B. Blum,[5] Himla Soodyall,[6] Mattias Jakobsson[1,7]*

The history of click-speaking Khoe-San, and African populations in general, remains poorly understood. We genotyped ~2.3 million single-nucleotide polymorphisms in 220 southern Africans and found that the Khoe-San diverged from other populations ≥100,000 years ago, but population structure within the Khoe-San dated back to about 35,000 years ago. Genetic variation in various sub-Saharan populations did not localize the origin of modern humans to a single geographic region within Africa; instead, it indicated a history of admixture and stratification. We found evidence of adaptation targeting muscle function and immune response; potential adaptive introgression of protection from ultraviolet light; and selection predating modern human diversification, involving skeletal and neurological development. These new findings illustrate the importance of African genomic diversity in understanding human evolutionary history.

Genetic, anthropological, and archaeological studies provide substantial support for an African origin of modern humans, but the process by which modern humans arose has been vigorously debated (*1*, *2*). African populations show the greatest genetic diversity, with genetic variation in Eurasia, Oceania and the Americas largely being a subset of the African diversity (*3*–*6*), with limited contribution from archaic humans (*7*). Within Africa, click-speaking southern African San and Khoe populations ["Khoe-San" from here on, following the San Council recommendations] harbor the deepest mitochondrial DNA lineages (*5*), have great genomic diversity (*8*–*10*), and probably represent the deepest historical population divergences among extant human populations (*11*, *12*). However, African populations have been underrepresented in genome-wide studies of genetic diversity, including assessment of the ethnic diversity within the Khoe-San in southern Africa, where previous studies have focused either on single-locus markers (*13*) or a few individuals from one or two populations (*3*, *4*, *8*–*10*).

We genotyped, quality-filtered, and phased ~2.3 million single-nucleotide polymorphisms (SNPs) in 220 individuals representing 11 populations from southern Africa: Ju/'hoansi, !Xun, /Gui and //Gana, Karretjie People (hereafter "Karretjie"), ≠Khomani, Nama, Khwe, "Coloured" (Colesberg), "Coloured" (Wellington), Herero, and Bantu-speakers (South Africa) [Fig. 1A, (*14*), and table S4]. These data were analyzed together with published data (*4*, *9*, *10*, *15*) after the removal of related and recently admixed individuals (*14*). To minimize the potential effect of ascertainment bias on results, we used several approaches that have previously been shown to be robust to these biases, including analyzing haplotypes, using minor allele frequency filtering within populations, and comparing results to available sequence data (*14*). In a principal components analysis (PCA), the first two PCs closely recapitulate many aspects of a geographic map of Africa [Fig. 1B, Procrustes correlation: 0.585, $P < 10^{-5}$ (*14*)], with the first PC representing a north-south axis that separates southern African Khoe-San populations from other populations, and the second PC representing an east-west axis that separates east African populations (including Hadza and Sandawe hunter-gatherers) from central African hunter-gatherers (Mbuti and Biaka Pygmies) and Niger-Kordofanian speakers (Fig. 1B). In this two-dimensional representation of sub-Saharan genetic diversity, hunter-gatherer populations from southern, central, and eastern Africa constitute three extremes, respectively, of a scaffold, where the fourth extreme is represented by all Niger-Kordofanian–speaking groups from across the African continent. Although Niger-Kordofanian–speaking populations have been sampled from southern, eastern, and western Africa, they all cluster closely in the vicinity of West African populations (Fig. 1B), a consequence of the recent "Bantu expansion." If Bantu-speaking populations are removed from the analysis, the correlation between the first two PCs and geography increases to 0.715 ($P < 10^{-5}$). In addition to geography, genetic structure can also be correlated with language and subsistence strategies, and we assessed the capacity of these factors to predict genetic components in sub-Saharan Africa (*14*). Geography predicted genetic components better than either language or subsistence, but combining geographic information with subsistence and especially linguistic information improved the prediction (Fig. 1E), suggesting that all of these factors contribute to genetic structure in sub-Saharan Africa.

Genetic cluster analysis (*16*) showed substantial structure among sub-Saharan individuals and reiterated the substructure among Khoe-San populations, Niger-Kordofanian speakers, east African populations, and central African hunter-gatherers (Fig. 2B) (*14*). Increasing the number of allowed clusters distinguishes finer levels of population substructure (Fig. 2B), including distinct non-African ancestry components for individuals who self-identify as "Coloured" (figs. S16, S18, and S21). Within the Khoe-San group, there was a distinct separation of Northern San populations (Ju speakers: !Xun and Ju/'hoansi) and

[1]Department of Evolutionary Biology, Uppsala University, Norbyvägen 18D, 752 36 Uppsala, Sweden. [2]Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA. [3]Department of Integrative Biology, University of California, Berkeley, CA 94720, USA. [4]Department of Anthropology and Archaeology, University of South Africa, Pretoria, South Africa. [5]Laboratoire TIMC-IMAG UMR 5525, Université Joseph Fourier, Centre National de la Recherche Scientifique, Grenoble, France. [6]Human Genomic Diversity and Disease Research Unit, Division of Human Genetics, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand and National Health Laboratory Service, Johannesburg, South Africa. [7]Science for Life Laboratory, Uppsala University, Uppsala, Sweden.

*To whom correspondence should be addressed. E-mail: carina.schlebusch@ebc.uu.se (C.M.S.); mattias.jakobsson@ebc.uu.se (M.J.)
†These authors contributed equally to this work.