

# SCIENTIFIC DATA

**OPEN**

**SUBJECT CATEGORIES**

- » Palaeoceanography
- » Biogeography
- » Palaeoecology

Received: 30 September 2015

Accepted: 18 November 2015

Published: 08 December 2015

## A global planktic foraminifer census data set for the Pliocene ocean

Harry Dowsett<sup>1</sup>, Marci Robinson<sup>1</sup> & Kevin Foley<sup>1</sup>

This article presents data derived by the USGS Pliocene Research, Interpretation and Synoptic Mapping (PRISM) Project. PRISM has generated planktic foraminifer census data from core sites and outcrops around the globe since 1988. These data form the basis of a number of paleoceanographic reconstructions focused on the mid-Piacenzian Warm Period (3.264 to 3.025 million years ago). Data are presented as counts of individuals within 64 taxonomic categories for each locality. We describe sample acquisition and processing, age dating, taxonomy and archival storage of material. These data provide a unique, stratigraphically focused opportunity to assess the effects of global warming on marine plankton.

<b>Design Type(s)</b>	observation design • biodiversity assessment objective
<b>Measurement Type(s)</b>	planktic foraminifera identification
<b>Technology Type(s)</b>	core sampling
<b>Factor Type(s)</b>	
<b>Sample Characteristic(s)</b>	Foraminifera • Bacillariophyta • North Atlantic Ocean • Pacific Ocean • Atlantic Ocean • South Atlantic Ocean • South Pacific Ocean • Indian Ocean • Southern Ocean • North America • New Zealand • continental shelf • ocean

<sup>1</sup>US Geological Survey, Eastern Geology and Paleoclimate Science Center, Reston, Virginia 20192, USA. Correspondence and requests for materials should be addressed to H.D. (email: hdowsett@usgs.gov).

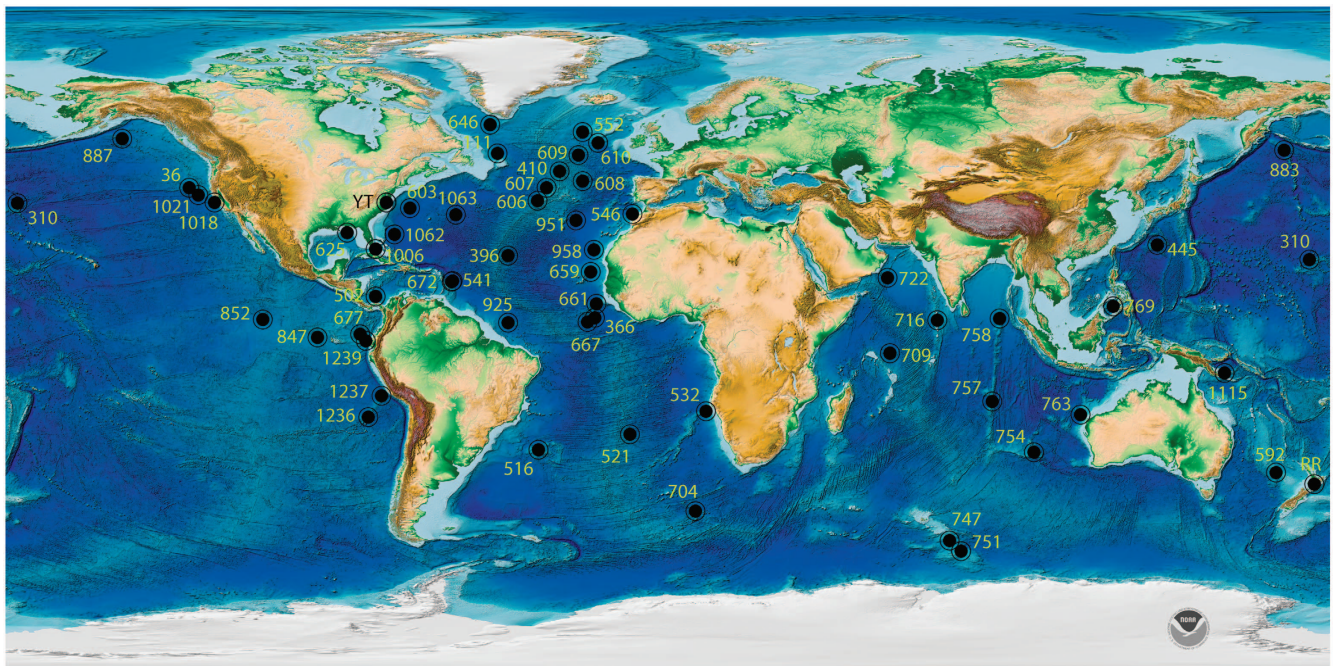
## Background and Summary

The Pliocene (5.3 to 2.6 million years ago (Ma)), specifically the mid-Piacenzian<sup>1,2</sup> (3.6 to 2.6 Ma), has been a focus of synoptic paleoclimate research for the past 25 years. The mid Piacenzian warm period (3.264 to 3.025 Ma) is the most recent time in Earth's past that exhibited climates not unlike those projected for the end of the 21st century<sup>3</sup>. With widespread recognition by most experts that anthropogenic drivers are *extremely likely* to have been the dominant cause of observed warming since the mid-20th century<sup>4</sup>, and surface temperatures projected to rise over the 21st century under all emission scenarios<sup>5</sup>, understanding the Pliocene climate has taken on new importance. While not a direct analog to future climate conditions, there is much to learn about the magnitude and spatial distribution of change from this, in essence, natural climate laboratory.

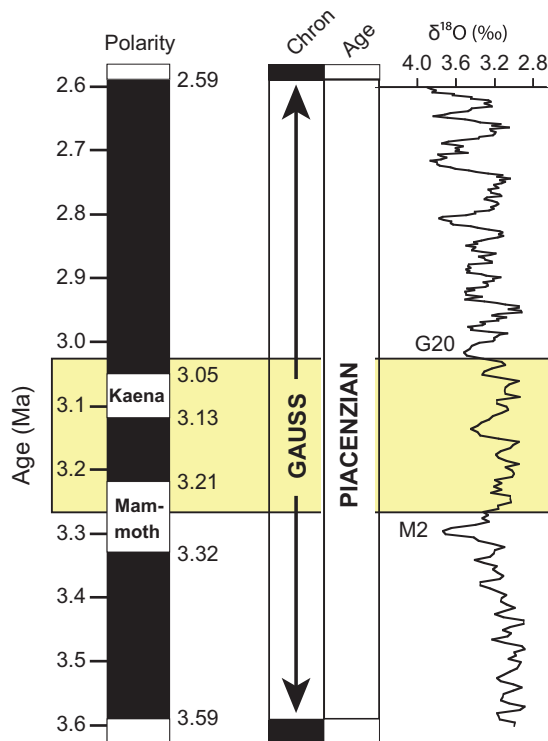
Since 1988 the United States Geological Survey (USGS) has developed a large-scale data collection project: PRISM (Pliocene Research, Interpretation and Synoptic Mapping)<sup>6</sup>. Over this time PRISM has produced a series of ever more complex global paleoenvironmental reconstructions that provide probable estimates of Piacenzian ocean temperatures, sea level, sea ice extent, land ice distribution, vegetation or land cover, and elevation<sup>6–11</sup>. PRISM is the most detailed global reconstruction of Earth conditions for a past period of global warmth. The PRISM reconstructions serve two purposes: (1) they provide a conceptual model of mid-Piacenzian conditions and (2) they are formatted for use as boundary condition data sets as well as verification data for climate models.

Various elements of the PRISM reconstruction have been used in climate modelling experiments to test hypotheses and assess the performance of the models<sup>9,11–15</sup>. The latest PRISM reconstructions have been used by a number of climate modelling groups in the Pliocene Model Intercomparison Project (PlioMIP)<sup>16</sup>. PRISM research has documented a reduced pole to equator surface temperature gradient in both marine and terrestrial settings<sup>13,17,18</sup>, reduced longitudinal temperature gradients in the equatorial Pacific<sup>19,20</sup>, reduced sea ice and changes in ocean circulation<sup>7,8</sup>, elevated sea levels<sup>21</sup> and major shifts in vegetation<sup>22–24</sup>.

While the PRISM reconstruction has terrestrial, marine and cryospheric components, the marine SST reconstruction has always been at the center of USGS PRISM work, and the faunal assemblage based SST data set has been the cornerstone of PRISM marine reconstructions. These SST reconstructions are based upon quantitative analysis of a large (>700,000 specimens) collection of mid-Piacenzian planktic foraminiferal data. Thus the PRISM planktic foraminifer collection, a census of individuals identified to species level from a global network of deep sea cores (Fig. 1) forms the basis for many of the PRISM paleoceanographic reconstructions<sup>25–40</sup> and have been used by others in ecological niche modelling<sup>41</sup> and analysis of diversity changes associated with global warming<sup>42</sup>. These data have been generated at the USGS since 1988, and additional data are being generated as part of the PRISM4 Paleoenvironmental reconstruction.



**Figure 1.** Map of sample locations from which Pliocene planktic foraminiferal census data were derived. Base map NOAA ETOPO1 global relief model<sup>54</sup>.



**Figure 2.** Chronostratigraphic framework for the mPWP. Position of the PRISM ‘time slab’ or mPWP (yellow shading) with respect to paleomagnetic stratigraphy (Gauss Chron) and the LR04 stack<sup>43</sup> of benthic  $\delta^{18}\text{O}$  records. The mPWP extends from 3.264 Ma (within the Mammoth subchron) to 3.025 Ma (just above Kaena subchron)<sup>8</sup> Marine Isotope stages G20 and M2 are identified. Black is normal polarity, white reversed polarity.

We present here raw faunal census data from 1,957 samples at 61 of our sites generated between 1988 and 2013 (Table 1 (available online only)). These data exist as counts of individual planktic forams placed into 64 taxonomic categories for each sample at each location. Samples are generally restricted to the Piacenzian Age as determined through a combination of magnetobiochronology and correlation of stratigraphic time series to the LR04 Marine Isotope Stages<sup>43</sup>.

## Methods

Our methodology for producing these data can be divided into three areas: chronology, sample acquisition and processing, and species identification.

### Chronology

All samples in the PRISM database are from the Pliocene Epoch and most fall within the Piacenzian Age. The PRISM ‘time slab’ or mid-Piacenzian Warm Period (mPWP) was originally defined as a 300 kyr interval of easily recognized warmth in the North Atlantic basin, centered on 3.0 Ma<sup>44</sup>. It was initially located in marine sections using magnetobiochronologic events. Over the past 25 years, events used to designate the mPWP have changed, and the geologic time scale used to calibrate those events has been refined and revised. Some newer sequences are dated by tuning stratigraphic records to known insolation changes caused by cyclical variations of the Earth’s orbit. The interval of time, the mPWP, the last time Earth experienced warming on the scale projected for the end of the 21st Century, has remained the same throughout the project.

The mPWP is presently defined as the period between the transition of marine isotope stages (MIS) M2/M1 (3.264 Ma) and G21/G20 (3.025 Ma) in the middle part of the Gauss Polarity Chron (Fig. 2)<sup>8</sup>. This interval ranges from C2An2r (Mammoth reversed polarity) to near the bottom of C2An1 (just above Kaena reversed polarity). This 239 kyr time slab correlates in part to planktonic foraminiferal zones PL3 (*Sphaeroidinellopsis seminulina* Highest Occurrence Zone), PL4 (*Dentoglobigerina altispira* Highest Occurrence Zone) and PL5 (Atlantic) (*Globorotalia miocenica* Highest Occurrence Zone) or PL5 (Indo-Pacific) (*Globorotalia pseudomiocenica* Highest Occurrence Zone)<sup>45</sup>.

Age determinations presented here are based on the best available data at the time of original investigation, however data contained in this archive were generated over a period of 25 years (Supplementary File 1). Age models for most core sites contained in the PRISM planktic foram census data set are based upon biochronology (calibrated first and last occurrence events for faunal and floral

taxa), magnetostratigraphy (dated paleomagnetic reversals), magnetobiochronology (combination of biochronologic and paleomagnetic calibrated events), tephrochronology (radiometrically dated ash beds), graphic correlation (Shaw's method of correlation via a magnetobiochronological model) or astronomical tuning (direct or indirect correlation of time series [usually  $\delta^{18}\text{O}$ ] to orbital forcing). Astronomical tuning was applied to these sequences: (607, 610, 659, 763, 806, 847, 852, 925, 1014, 1237, and 1239).

Due to the inconsistencies of calibrated datums both regionally and over the time period these cores were analyzed, as well as the many versions of geological time scales, users are urged to research and develop their own age models for these sites. Comparison of samples from one location to another, based upon provided ages, will result in diachronous correlations. Therefore, users are advised to consult the most current paleontological and chronological data for these sites.

### Sample acquisition and processing

The majority of samples come from cores raised by DSDP, ODP, and from a number of outcrops on land. For marine cores, 10–20 cc samples were removed from a split core using a cylindrical plug and sealed in a plastic bag for transport to the USGS. Outcrop samples were retrieved using a hand shovel and rock hammer to obtain approximately 50–500 g of sediment. Samples were placed in plastic bags for transport to the USGS. In the lab, samples were oven dried at  $\leq 50^\circ\text{C}$ , and then soaked and agitated in water with  $\sim 2$  ml of dilute sodium hexametaphosphate solution (5 g to 1 l water) for 1–2 h. Samples were then washed over a 63  $\mu$  or 150  $\mu$  sieve until clean. Samples were then oven dried at  $\leq 50^\circ\text{C}$ , then dry-sieved to concentrate the  $\geq 150 \mu$  fraction. The  $\geq 150 \mu$  fraction was placed in a sample splitter and split until  $\sim 300$  planktic foraminifera specimens were obtained. There is a 0.05 probability that we failed to detect a taxon represented by 3 individuals (1%) in a population of 300 individuals<sup>46</sup>. Reducing the probability to 0.01 would require counting an additional 200 specimens. Census counts are labor-intensive, and using 300 specimens is common practice in studies similar to ours. In samples that did not contain 300 planktic foraminifer specimens, all planktic foraminifers were counted. Specimens were placed on a Plummer slide (60 cell faunal micro slide) for identification and sorting into 64 possible taxonomic categories (Supplementary File 2). Foraminifers were manipulated with a fine (00000) paintbrush under an incident light microscope and fixed to the slide using a weak, water-soluble glue.

### Species identification and archival samples

Individual specimens were identified to species level following taxonomic concepts of Parker<sup>47,48</sup>, Blow<sup>49</sup>, and Dowsett and Robinson<sup>30</sup> (Supplementary File 2). All counts were generated by the PRISM Project; Data published by others are not included in this release.

Foraminifers were grouped by species and fixed in place on slides, and additional washed residue (when available) for all samples shown in the global planktic foraminifer census database, are physically archived at the US Geological Survey in Reston, Virginia, USA.

### Data Records

The census of 593,676 individuals identified to species level in 1,957 Pliocene age ocean sediment samples is accessible at National Climate Data Center (NCDC) (Data Citation 1: Global Planktic Foraminifer Census Data Set for the Pliocene Ocean <https://www.ncdc.noaa.gov/paleo/study/19281>). The data for each sample consist of location information (name and geographic coordinates), sample number, position in stratigraphic sequence (depth below sea floor for sediment cores and height above base of land section for terrestrial outcrops), age, and number of individuals assigned to each of 64 taxonomic categories (see Supplementary Files 1 and 2).

### Technical Validation

Micropaleontological techniques for processing and sorting individual foraminifer tests into species are well documented and standardized in the paleoceanographic community<sup>50–52</sup>. Large projects, similar to PRISM, have generated planktic foraminiferal census data and an important factor for these studies has been maintaining internal consistency in identifications<sup>53</sup>. We maintain consistency and avoid variation in identification of species by having a small number of individuals with the same taxonomic concepts identify all specimens. We further reduce the possibility of taxonomic drift by having all identifications checked by one micropaleontologist associated with the project since its inception. We consider this taxonomic consistency a primary strength of our data.

### References

1. Castradori, D., Rio, D., Hilgen, F. J. & Lourens, L. J. The global standard stratotype-section and point (GSSP) of the Piacenzian (Middle Pliocene). *Episodes* **21**, 88–93 (1998).
2. Gibbard, P. L., Head, M. J. & Walker, M. J. C. The Subcommission on Quaternary Stratigraphy. Formal ratification of the Quaternary System/Period and the Pleistocene Series/Epoch with a base at 2.58 Ma. *Journal of Quaternary Science* **25**, 96–102 (2010).
3. IPCC. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, 2013).
4. Verheggen, B. et al. Scientists' Views about Attribution of Global Warming. *Environmental Science & Technology* **48**, 8963–8971 (2014).

5. IPCC. *Working Group I Contribution to the IPCC Fifth Assessment Report (AR5), Climate Change 2013: The Physical Science Basis* (Intergovernmental Panel on Climate Change, 2013).
6. Dowsett, H. J. in *Deep-time perspectives on climate change: marrying the signal from computer models and biological proxies* (eds Williams M., Haywood A. M., Gregory J. & Schmidt D. N.) 459–480 (Micropalaeontological Society (Special Publication), Geological Society of London, 2007).
7. Dowsett, H., Robinson, M. & Foley, K. Pliocene three-dimensional global ocean temperature reconstruction. *Climate of the Past* **5**, 769–783 (2009).
8. Dowsett, H. *et al.* The PRISM3D paleoenvironmental reconstruction. *Stratigraphy* **7**, 123–139 (2010).
9. Dowsett, H. J. *et al.* Middle Pliocene Paleoenvironmental Reconstruction: PRISM 2. *U.S. Geological Survey, Open File Report 99-535*, available online only <http://pubs.usgs.gov/of/1999/of99-535/> (1999).
10. Pound, M. J. *et al.* Late Pliocene lakes and soils: a global data set for the analysis of climate feedbacks in a warmer world. *Climate of the Past* **10**, 167–180 (2014).
11. Sohl, L. E. *et al.* PRISM3/GISS topographic reconstruction. *Geological Survey Data Series* **419**, available online only <http://pubs.usgs.gov/ds/419/> (2009).
12. Chandler, M. A., Rind, D. & Thompson, R. Joint investigations of the middle Pliocene climate II: GISS GCM Northern Hemisphere results. *Global and Planetary Change* **9**, 197–219 (1994).
13. Dowsett, H. J. *et al.* Sea Surface Temperature of the mid-Piacenzian Ocean: A Data-Model Comparison. *Sci. Rep.* **3**, 1–8 (2013).
14. Lunt, D. J. *et al.* On the causes of mid-Pliocene warmth and polar amplification. *Earth and Planetary Science Letters* **321–322**, 128–138 (2012).
15. Zhang, Z. S. *et al.* Mid-pliocene Atlantic Meridional Overturning Circulation not unlike modern. *Clim. Past* **9**, 1495–1504 (2013).
16. Haywood, A. *et al.* Large-scale features of Pliocene climate: results from the Pliocene Model Intercomparison Project. *Climate of the Past* **9**, 191–209 (2013).
17. Dowsett, H. J. *et al.* Micropaleontological Evidence for Increased Meridional Heat Transport in the North Atlantic Ocean During the Pliocene. *Science* **258**, 1133–1135 (1992).
18. Salzmann, U. *et al.* Challenges in quantifying Pliocene terrestrial warming revealed by data-model discord. *Nature Climate Change* **3**, 969–974 (2013).
19. Dowsett, H. J. Faunal re-evaluation of Mid-Pliocene conditions in the western equatorial Pacific. *Micropaleontology* **53**, 447–456 (2007).
20. Dowsett, H. J. & Robinson, M. M. Mid-Pliocene equatorial Pacific sea surface temperature reconstruction: a multi-proxy perspective. *Philosophical Transactions of the Royal Society, A* **367**, 109–125 (2009).
21. Dowsett, H. J. & Cronin, T. M. High eustatic sea level during the middle Pliocene: Evidence from the southeastern U.S. Atlantic Coastal Plain. *Geology* **18**, 435–438 (1990).
22. Salzmann, U., Haywood, A. M. & Lunt, D. J. The past is a guide to the future? Comparing Middle Pliocene vegetation with predicted biome distributions for the twenty-first century. *Philosophical Transactions of the Royal Society, A* **367**, 189–204 (2009).
23. Salzmann, U., Haywood, A. M., Lunt, D. J., Valdes, P. J. & Hill, D. J. A new global biome reconstruction and data-model comparison for the Middle Pliocene. *Global Ecology and Biogeography* **17**, 432–447 (2008).
24. Thompson, R. S. & Fleming, R. F. Middle Pliocene vegetation: reconstructions, paleoclimatic inferences, and boundary conditions for climate modeling. *Marine Micropaleontology* **27**, 27–49 (1996).
25. Caballero, R. & Dowsett, H. J. Mid-Pliocene planktic foraminifer census data from Ocean Drilling Program Hole 1237C. *U.S. Geological Survey Data Series* **336**, available online only <http://pubs.usgs.gov/ds/336> (2008).
26. Dowsett, H. J. & Caballero, R. Mid-Pliocene planktic foraminifer census data from Ocean Drilling Program Hole 847C. *U.S. Geological Survey Data Series*, **295**, available online only <http://pubs.usgs.gov/ds/295> (2007).
27. Dowsett, H. J., Gosnell, L. B. & Poore, R. Z. Pliocene planktic foraminifer census data from Deep Sea Drilling Project Holes 366A, 410, 606, and 646B. *U.S. Geological Survey, Open File Report 88-654*, available online only <http://pubs.er.usgs.gov/publication/ofr88654> (1988).
28. Dowsett, H. J., Ishman, S. E., Verardo, S. & Polanco, E. Pliocene planktic foraminifer census data from Ocean Drilling Program Hole 704A. *U.S. Geological Survey Open File Report 93-304*, available online only <http://pubs.er.usgs.gov/publication/ofr92418> (1993).
29. Dowsett, H. J. & Polanco, E. F. Pliocene planktic foraminifer census data from Deep Sea Drilling Project Holes 541 and 546. *U.S. Geological Survey Open File Report 92-418*, available online only <http://pubs.er.usgs.gov/publication/ofr92418> (1992).
30. Dowsett, H. J. & Robinson, M. M. Mid-Pliocene planktic foraminifer assemblage of the North Atlantic Ocean. *Micropaleontology* **53**, 105–126 (2007).
31. Dowsett, H. J. & West, S. Pliocene planktic foraminifer census data from Deep Sea Drilling Project Hole 445. *U.S. Geological Survey Open File Report 93-307*, available online only <http://pubs.er.usgs.gov/publication/ofr93307> (1993).
32. Dowsett, H. J. & West, S. M. Pliocene planktic foraminifer census data from Deep Sea Drilling Project Hole 607 and Ocean Drilling Program Hole 661A. *U.S. Geological Survey, Open File Report 92-413*, available online only <http://pubs.er.usgs.gov/publication/ofr92413> (1992).
33. Dowsett, H. J. & Wiggs, L. B. Planktonic foraminiferal assemblage of the Yorktown Formation, Virginia, USA. *Micropaleontology* **38**, 75–86 (1992).
34. Foley, K. M. & Dowsett, H. J. Pliocene planktic foraminifer census data from Deep Sea Drilling Project Holes 667 and 659A. *U.S. Geological Survey, Open File Report 92-434*, available online only <http://pubs.er.usgs.gov/publication/ofr92434> (1992).
35. Polanco, E. & Dowsett, H. Pliocene planktic foraminifer census data from Deep Sea Drilling Project Hole 463 and Ocean Drilling Program Hole 769B. *U.S. Geological Survey Open-File Report 93-308*, available online only <http://pubs.er.usgs.gov/publication/ofr93308> (1993).
36. PRISM. Pliocene planktic foraminifer census data from the North Atlantic region. *U.S. Geological Survey Open-File Report 96-669*, available online only <http://pubs.er.usgs.gov/publication/ofr96669> (1996).
37. Robinson, M. M. *et al.* Mid-Pliocene planktic foraminifer census data and alkenone unsaturation indices from Ocean Drilling Hole 677A. *U.S. Geological Survey Data Series* **353**, available online only <http://pubs.usgs.gov/ds/353> (2008).
38. Robinson, M. M. & Dowsett, H. J. Pliocene planktic foraminifer census data from DSDP Site 592, southwest Pacific Ocean. *U.S. Geological Survey, Open File Report 96-544*, available online only <http://pubs.er.usgs.gov/publication/ofr96544> (1996).
39. Wiggs, L. B. & Dowsett, H. J. Pliocene planktic foraminifer census data from Deep Sea Drilling Project Hole 396 and Ocean Drilling Program Hole 672. *U.S. Geological Survey, Open File Report 92-414*, available online only <http://pubs.er.usgs.gov/publication/ofr92414> (1992).
40. Wiggs, L. B. & Poore, R. Z. Pliocene planktic foraminifer census data from Deep Sea Drilling Project Holes 502A, B, C. *U.S. Geological Survey, Open File Report 91-325*, available online only <http://pubs.er.usgs.gov/publication/ofr91325> (1991).
41. Saupe, E. E. *et al.* Macroevolutionary consequences of profound climate change on niche evolution in marine molluscs over the past three million years. *Proceedings of the Royal Society B* **281**, 2014995 (2014).
42. Yasuhara, M., Hunt, G., Dowsett, H. J., Robinson, M. M. & Stoll, D. K. Latitudinal species diversity gradient of marine zooplankton for the last three million years. *Ecology Letters* **15**, 1174–1179 (2012).

43. Lisiecki, L. E. & Raymo, M. E. A Pliocene-Pleistocene stack of 57 globally distributed benthic  $\delta^{18}\text{O}$  records. *Paleoceanography* **20**, PA1003 (2005).
44. Dowsett, H. & Poore, R. Pliocene sea surface temperatures of the North Atlantic Ocean at 3.0 Ma. *Quaternary Science Reviews* **10**, 189–204 (1991).
45. Wade, B. S., Pearson, P. N., Berggren, W. A. & Pälike, H. Review and revision of Cenozoic tropical planktonic foraminiferal biostratigraphy and calibration to the geomagnetic polarity and astronomical time scale. *Earth Science Reviews* **104**, 111–142 (2011).
46. Hay, W. W. Probabilistic Stratigraphy. *Eclogae Geologicae Helveticae* **65**, 255–266 (1972).
47. Parker, F. L. Planktonic foraminiferal species in Pacific sediments. *Micropaleontology* **8**, 219–254 (1962).
48. Parker, F. L. Late Tertiary biostratigraphy (planktonic foraminifera) of tropical Indo-Pacific deep-sea cores. *Bulletin of American Paleontology* **52**, 115–208 (1967).
49. Blow, W. H. in *Proceedings of the 1st International Conference on Planktonic Microfossils* (eds Bronniman P. & Renz H. H.) 199–422 (Brill, E. J., 1969).
50. Armstrong, H. A. & Brasier, M. D. *Microfossils*. Second edn (Blackwell Publishing, 2005).
51. Haynes, J. R. *Foraminifera* (John Wiley & Sons, 1981).
52. Imbrie, J., Kipp, N. G. in *The Late Cenozoic Glacial Ages* (ed Turekian K. K.) 71–181 (Yale University Press, 1971).
53. Cline, R. M. & Hays, J. D. in *Memoir 145* (Geological Society of America, 1976).
54. Amante, C. & Eakins, B. W. *ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis*. Vol. NOAA Technical Memorandum NESDIS NGDC-24 (National Geophysical Data Center, NOAA, 2009).

## Data Citation

1. Foley, K., Dowsett, H. & Robinson, M. *National Climate Data Center* <https://www.ncdc.noaa.gov/paleo/study/19281> (2015).

## Acknowledgements

We thank the PRISM researchers who contributed to data collection: Brendan Lutz, Dick Poore, Danielle Stoll, Emerson Polanco, Linda Wiggs, Rocio Caballero-Gill, Stacy Verardo, Stephanie Strother, Stephanie West, and Sarah Brown. Samples used in this study were provided by the International Ocean Discovery Program. We acknowledge the continued support of the USGS Climate and Land Use Research and Development Program. This is a product of the PRISM Project.

## Author Contributions

H.D. drafted the manuscript and undertook data collection. K.F. compiled dataset and undertook data collection. M.R. undertook data collection. All authors read and approved the final manuscript.

## Additional Information

Table 1 is only available in the online version of this paper.

Supplementary Information accompanies this paper at <http://www.nature.com/sdata>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Dowsett, H. *et al.* A global planktic foraminifer census data set for the Pliocene ocean. *Sci. Data* **2**:150076 doi: 10.1038/sdata.2015.76 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0>

Metadata associated with this Data Descriptor is available at <http://www.nature.com/sdata/> and is released under the CC0 waiver to maximize reuse.