

Post-invasion demography of prehistoric humans in South America

Amy Goldberg^{1*}, Alexis M. Mychajliw^{1*} & Elizabeth A. Hadly^{1,2}

As the last habitable continent colonized by humans, the site of multiple domestication hotspots, and the location of the largest Pleistocene megafaunal extinction, South America is central to human prehistory^{1–7}. Yet remarkably little is known about human population dynamics during colonization, subsequent expansions, and domestication^{2–5}. Here we reconstruct the spatiotemporal patterns of human population growth in South America using a newly aggregated database of 1,147 archaeological sites and 5,464 calibrated radiocarbon dates spanning fourteen thousand to two thousand years ago (ka). We demonstrate that, rather than a steady exponential expansion, the demographic history of South Americans is characterized by two distinct phases. First, humans spread rapidly throughout the continent, but remained at low population sizes for 8,000 years, including a 4,000-year period of ‘boom-and-bust’ oscillations with no net growth. Supplementation of hunting with domesticated crops and animals^{4,8} had a minimal impact on population carrying capacity. Only with widespread sedentism, beginning ~ 5 ka^{4,8}, did a second demographic phase begin, with evidence for exponential population growth in cultural hotspots, characteristic of the Neolithic transition worldwide⁹. The unique extent of humanity’s ability to modify its environment to markedly increase carrying capacity in South America is therefore an unexpectedly recent phenomenon.

Genetic, archaeological and linguistic evidence suggest the first Americans descended from ancestral Siberians. A small population or populations crossed the land bridge connecting Asia and Alaska between 15 ka and 30 ka, reaching southern South America by at least 14.5 ka^{1–6}. Yet the peopling of South America and the relative effects of climate and culture on early and mid-Holocene population dynamics remains unclear^{2–5,8,10,11}, particularly during the rise of agriculture worldwide.

More generally, with its recent and rapid colonization, South America provides a unique opportunity to study the colonization behaviour and population growth dynamics of modern humans. South America is exceptional as it was peopled during a single wave of invasion over a narrow time window, with limited later migration until historic times^{5,6}. Therefore, it provides a continental view of human population history unavailable elsewhere.

To discern the dynamics of the South American human populations through the Holocene, we compiled a database of 5,464 radiocarbon (¹⁴C) dates from 1,147 archaeological sites associated with human occupation (Fig. 1, Supplementary Data 1). We confined our analyses to dates ranging from late-Pleistocene colonization to 2 ¹⁴C ka. Dates are calibrated unless specified as radiocarbon dates, ¹⁴C ka. We use two proxies for human population size: the probability density of summed calibrated radiocarbon dates (SCPDs)^{12–15}, and the number of occupied archaeological sites over time^{13–16}. After applying quality control filters and pruning the data to prevent oversampling, we are left with a set of 2,576 approximately independent merged occupation events for SCPD analyses.

Archaeological evidence of the initial peopling of South America is scarce, with many of the earliest sites found in Patagonia, despite geographic constraints and genetic data suggesting a north-to-south colonization route^{1–6}. Indeed, once there is evidence of occupation, humans were already dispersed throughout the continent. Although humans are geographically widespread, initial site and date densities suggest low overall population sizes (Fig. 2a and Extended Data Figs 1 and 2). However, because sea levels were ~ 100 m lower during the interval of colonization and a coastal route of migration has been hypothesized, the earliest archaeological sites may be on the continental shelf that is now underwater^{3–6}.

Small human population sizes are almost immediately followed by a rapid increase in the density of radiocarbon dates and the number of occupied sites from 13 ka to 9 ka (Figs 2 and 3). At ~ 9 ka, the SCPD stabilizes, oscillating around a constant mean for $\sim 4,000$ years (Fig. 3). During this period, we find a peak in the SCPD occurring at ~ 11 ka, as well as recurring peaks from 9 ka to 5.5 ka.

It is tempting to interpret the initial peak around 11 ka and associated abrupt decline as evidence of overshooting carrying capacity after rapid colonization and correction to a carrying capacity over time. Such an overshoot is typical of species growing in a new environment, and is often linked to the rapid demise of a main food resource, perhaps indicative of the South American megafaunal extinction in this case^{7,17}. Notably, our simulations suggest the recurring mid-Holocene peaks and troughs (9 ka to 5.5 ka) cannot be explained by calibration artefacts alone, but rather represent a dynamic equilibrium with oscillation around a carrying capacity¹⁸ (Extended Data Figs 3 and 4). After 4,000

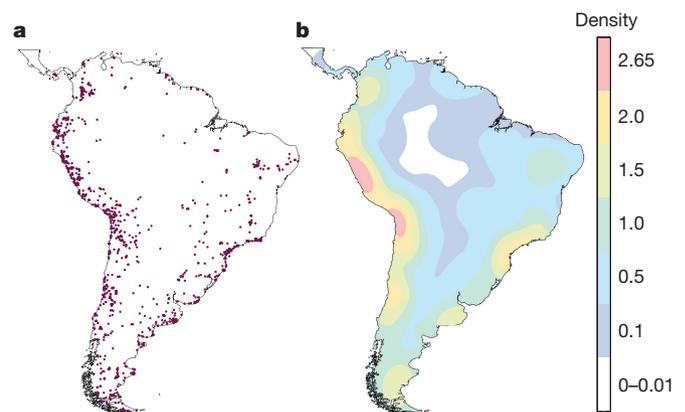


Figure 1 | Archaeological sites in database. **a**, Location of sites with radiocarbon dates with both laboratory numbers and latitude and longitude coordinates ($N=1,147$). **b**, Kernel-smoothed sampling intensity for the number of sampled sites used in analyses, with a search radius of 660 km. Density is measured by square decimal degree. Plotted in ArcGIS using the South America Albers Equal Area Conic Projection (standard parallels: $-5, -42$ degrees).

¹Biology Department, Stanford University, Stanford, California 94305, USA. ²Woods Institute, Stanford University, Stanford, California 94305, USA.

*These authors contributed equally to this work.

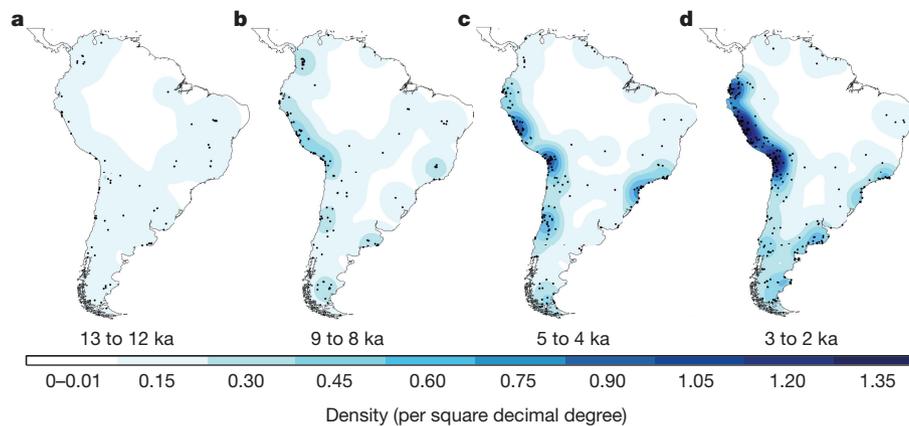


Figure 2 | Kernel density maps of occupied area. a–d, For 13 ka to 12 ka (a), 9 ka to 8 ka (b), 5 ka to 4 ka (c), and 3 ka to 2 ka (d), the kernel density smoothing of occupied area based on sites occupied in 1,000-year bins, with site locations, plotted in ArcGIS.

years of a dynamic equilibrium, there was renewed population growth, which was not followed by levelling off in our data set.

Kernel density maps of the land occupied by humans from ~14 ka to 2 ka suggest clear patterns in known centres of growth, with the earliest and highest population growth patterns seen on the Pacific coastline in Peru, Chile and Ecuador. To a lesser extent, we also witnessed centres of growth in Patagonia and the Brazilian Coastal Strip during the later mid-Holocene (Fig. 2 and Extended Data Figs 1 and 2). Lack of signal in certain regions, such as the Amazon, may reflect poor preservation of material or a lack of archaeological sampling, rather than an absence of people (Fig. 1b). South American population size dynamics show substantial spatial and temporal heterogeneity, which must be accounted for in models of human biological and cultural evolution.

For a first approximation of the overall growth in population size, we compared the mean density of the SCPD from 13.5 ka to 2.5 ka. We estimate that population size increased by roughly 1,000-fold during this time. To quantitatively analyse population growth rates and trends, we considered three general models of long-term human population growth: (1) exponential, (2) logistic, and (3) climate-mediated (Extended Data Table 1). Human populations are frequently modelled using exponential growth, uninhibited growth at a constant rate over time^{12,19,20}. Recent discussion has focused around the explosive growth of the last few hundred years of global human populations^{20–22}, but the applicability of an exponential growth model for the majority of human history remains unclear. An alternative to unconstrained growth is density-dependent or logistic growth. As populations expand into new and favourable habitats, they often experienced rapid growth followed by a gradual decline in growth rate as they used resources and approached carrying capacity. This logistic growth model has been demonstrated theoretically and empirically in a variety of mammalian, avian, plant and bacterial populations^{22–25}. A final model for population dynamics—climate-mediated growth—is one where the resources are determined by top-down controls characteristic of fluctuating environments^{10,26}.

We found trajectories of exponential growth and/or climate-mediated growth to be poor fits for the long-term human population size in South America (Fig. 4 and Extended Data Table 1). Rather, under a likelihood-style framework, the best fit was a two-phase model with density-dependent growth from initial peopling through 5.5 ka, followed by a recent phase of exponential growth from 5.5 ka to 2 ka (Fig. 4; minimum $\Delta\text{BIC} = 11.7$, Schwarz weight >0.99). This model of growth is consistent with our kernel density analyses, and is distinct from the curvilinear shapes seen in SCPDs from North America, Europe, and Australia^{12,14,15}.

Under the model, initial growth was rapid ($r = 0.131\%$ per year, 3.33 to 3.74% per generation), with the population doubling every 19 to 21 generations (for a generation time between 25 to 28 years), followed by a relatively constant population size from 9 ka to 5.5 ka. Early growth

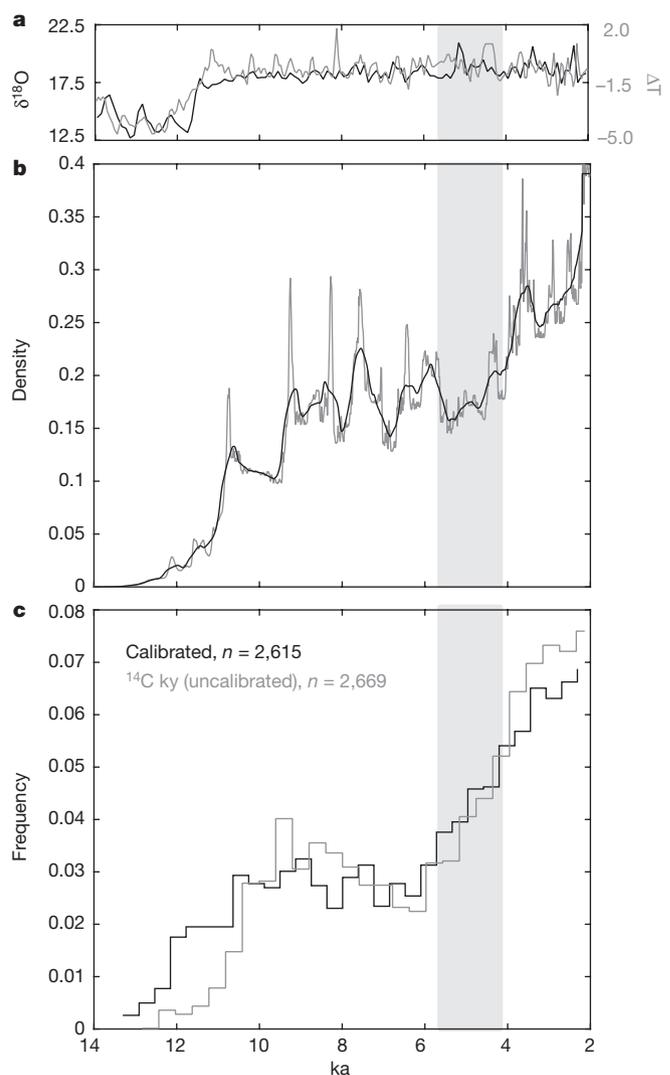


Figure 3 | Climatic and radiocarbon records for 14 ka to 2 ka. a, The 100-year averages of oxygen isotope data, $\delta^{18}\text{O}$ (‰), from Sajama, Bolivia ice core (black) and the difference in temperature from recent time average reconstructed from Vostok, Antarctica ice core (grey). b, SCPD for South American continent (grey), with a 400-year moving average (black). c, Frequency of occupied sites over time in 400-year time bins using the median of calibrated dates (95.4% distribution, black) and the frequency of occupied sites over time in 400-year time bins using uncalibrated radiocarbon dates (grey). For all panels, the grey bar indicates estimated time of demographic phase shift to renewed growth.

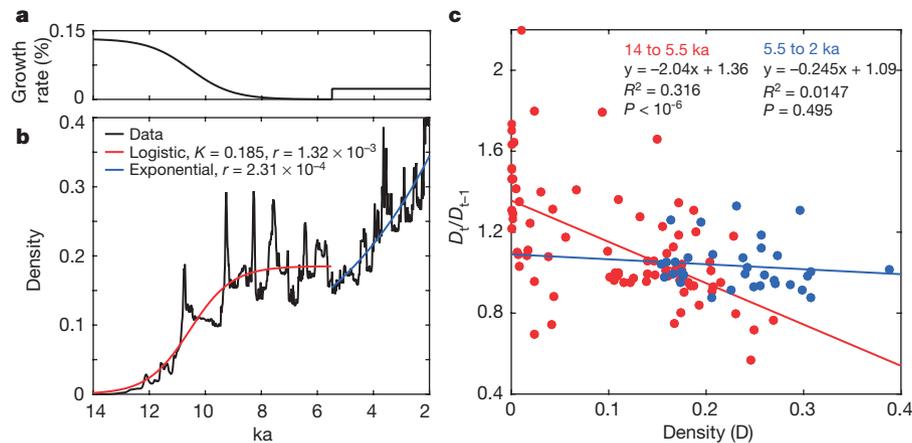


Figure 4 | A two-phase model for the human invasion of South America. **a**, Yearly growth rate (percentage) of the population over time under the model. **b**, SCPD for South American continent, with the best-fit model of population size over time: logistic growth (red) from colonization to 5.5 ka, and exponential growth from 5.5 ka to 2 ka. **c**, A linear model for density-dependent growth from 14 ka to 5.5 ka (red), with the rate of change of

density significantly negatively correlated with current density, and from 5.5 ka to 2 ka (blue), where there is no evidence for density-dependent growth. Rather, for 5.5 ka to 2 ka, the constant rate is characteristic of exponential growth. The density and rate of change of density in the SCPD is calculated in 100-year bins, with each point representing the mean density of that point \pm 50 years.

rates for South America are approximately an order of magnitude higher than archaeological and genetic estimates from the same period in Eurasia^{18,20,27}, and are on par with or larger than genetic estimates of explosive growth during the past 140 generations²¹.

Before the mid-Holocene, human colonization of South America resembled the population size dynamics of a typical invasive species^{24,25}, with rapid initial expansion and resource-limited growth over time. However, after thousands of years with a dynamic equilibrium, South American populations experienced renewed growth, indicating an increase in carrying capacity. This increase is a sharp departure from animal models of population growth because it does not correlate with an amelioration of climate or known environmental change favourable to humans (Fig. 3). The growth rate in the second phase is estimated as $r = 0.579$ to 0.649% per generation.

We compared expansion dynamics of human populations in South America to a general density-dependence model for invasive species by comparing the density of radiocarbon dates to the rate of change of radiocarbon dates adapted from^{24,25}. Fitting a linear model, the rate of population growth was inversely proportional to the current population size. We found a significant negative relationship for the period of inferred logistic growth. Conversely, the relationship was not significant during the period from 5.5 ka to 2 ka, consistent with exponential growth (Fig. 4c). The differing behaviour of the two linear models is additional evidence for two distinct phases of population dynamics.

Under our model, estimates for the yearly growth rate over time (Fig. 4a) can be used to infer population size (Extended Data Fig. 5). For an initial size of one thousand, this estimate of a \sim 600-fold increase is lower than, but consistent with, that from the ratio of radiocarbon dates at the start and end time periods alone. Together they give a range of approximately 615,000 to 1,000,000 people in South America by 2 ka, with more than half of the total population growth occurring between 5.5 ka to 2 ka.

Renewed growth is evidence of a change in carrying capacity in South America, with two, possibly interacting, causes: (1) climatic change, or (2) cultural or technological change. Palaeoclimate proxies do not support the hypothesis that the renewed growth was triggered by climatic change. Although the initial warming period (\sim 12–11 ka) corresponds to an increase in population size during colonization, oxygen isotope levels fluctuate around constant mean through most of the Holocene (Fig. 3a). Palaeoclimatic and vegetation studies suggest climatic change during \sim 6 ka to 4 ka is region-specific and poorly understood, with no known global mechanisms for change^{4,8,28}.

Many of the most culturally developed and largest populations of the mid-Holocene overlap with harsh, dry climates, such as the coastal Chinchorro culture, and the northern Las Vegas culture^{4,8} (Fig. 2 and Extended Data Fig. 6). Conversely, in Patagonia the population growth is more modest, consistent with lower levels of sedentism and domestication through the later Holocene^{4,8}.

South American population density during the mid-Holocene is in the lower range of estimates for the density of worldwide hunter-gatherer populations²⁷. As populations grew in the mid-Holocene, formally isolated populations may have interacted more and, through trade and the spread of innovations, acted as a buffer towards the environment increasing the carrying capacity. This pattern can be cyclical, with larger populations leading to more opportunity for innovation and, in turn, the capacity for even larger populations, as well as unique interactions between culture and local environments^{11,29}. We estimate population density increased threefold from 5.5 ka to 2 ka. The continental estimates for population density are low for agricultural populations²⁷, perhaps owing to regional variation in levels of agriculture^{4,8}.

South America had no known single cohesive culture comparable to Clovis or extensive evidence for long-range material and cultural sharing, likely due to divergent environments, geographic barriers to gene flow, and low population density^{4,8}. As such, domestication was slow, involving long periods with a handful of domesticated plants and few animals supplementing hunting and gathering, starting in the northwest approximately 9 ka, with multiple domestication hotspots across the continent. This pattern is a sharp contrast to the European Neolithic, characterized by a somewhat constant rate of agricultural expansion from a singular Middle Eastern origin^{9,30}. Even with vastly different domestication patterns and histories, South America surprisingly shows the same demographic transition characteristic of the European Neolithic⁹.

Despite the long history of supplementary domestication, including staple crops such as squash, peppers and maize^{4,8}, the renewed growth did not begin until the shift to a predominantly sedentary and agricultural subsistence. This process occurred throughout the continent between 5.5 ka and 3.5 ka, though it did not fully reach Patagonia, as this region likely then, as now, was not amenable to agriculture. The timing of this transition overlaps well with our estimated time of renewed population expansion (5.7 ka to 4.1 ka). Stable food sources, the need for human labourers, coupled with increased parental availability lead to shorter birth intervals and larger populations in agricultural societies⁹.

Humans are unique in the extent of their ability to manipulate their environments, and thus change local carrying capacity. Yet, we demonstrate that human prehistoric growth behaviour in fact resembles that of most animal populations—until the rise of agriculture. That is, although humans have been manipulating their environments through tool use and supplementary domesticated crops for millennia, it was not until the rise of widespread sedentism and agriculture that humans gained our distinctive ability to increase local and global carrying capacity.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 3 November 2015; accepted 26 January 2016.

Published online 6 April 2016.

- Meltzer, D. J. *First Peoples in a New World: Colonizing Ice Age America* (Univ. California Press, 2009).
- Raghavan, M. *et al.* Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science* **349**, aab3884 (2015).
- Dillehay, T. D. Probing deeper into first American studies. *Proc. Natl Acad. Sci. USA* **106**, 971–978 (2009).
- Moore, J. D. *A Prehistory of South America: Ancient Cultural Diversity on the Least Known Continent* (Univ. Press of Colorado, 2014).
- Rothhammer, F. & Dillehay, T. D. The late Pleistocene colonization of South America: an interdisciplinary perspective. *Ann. Hum. Genet.* **73**, 540–549 (2009).
- Reich, D. *et al.* Reconstructing Native American population history. *Nature* **488**, 370–374 (2012).
- Barnosky, A. D., Koch, P. L., Feranec, R. S., Wing, S. L. & Shabel, A. B. Assessing the causes of Late Pleistocene extinctions on the continents. *Science* **306**, 70–75 (2004).
- Silverman, H. & Isbell, W. *Handbook of South American Archaeology* (Springer, 2008).
- Bocquet-Appel, J. P. & Bar-Yosef, O. (eds) *The Neolithic Demographic Transition and its Consequences* (Springer, 2008).
- Sandweiss, D. H., Solís, R. S., Moseley, M. E., Keefer, D. K. & Ortloff, C. R. Environmental change and economic development in coastal Peru between 5,800 and 3,600 years ago. *Proc. Natl Acad. Sci. USA* **106**, 1359–1363 (2009).
- Marquet, P. A., Santoro, C. M., Latorre, C., Standen, V. G. & Abades, S. R. Emergence of social complexity among coastal hunter-gatherers in the Atacama Desert of northern Chile. *Proc. Natl Acad. Sci. USA* **109**, 14754–14760 (2012).
- Shennan, S. *et al.* Regional population collapse followed initial agriculture booms in mid-Holocene Europe. *Nature Commun.* **4**, 2486 (2013).
- Rick, J. W. Dates as data: an examination of the Peruvian preceramic radiocarbon record. *Am. Antiq.* **52**, 55–73 (1987).
- Williams, A. N. The use of summed radiocarbon probability distributions in archaeology: a review of methods. *J. Archaeol. Sci.* **39**, 578–589 (2012).
- Peros, M. C., Munoz, S. E., Gajewski, K. & Viau, A. E. Prehistoric demography of North America inferred from radiocarbon data. *J. Archaeol. Sci.* **37**, 656–664 (2010).
- Mellars, P. & French, J. S. Tenfold population increase in Western Europe at the Neandertal-to-modern human transition. *Science* **333**, 623–627 (2011).
- Barnosky, A. D. *et al.* The variable impact of Late-Quaternary megafaunal extinction in causing ecological state shifts in North and South America. *Proc. Natl Acad. Sci. USA* **113**, 856–861 (2016).
- Boone, J. L. Subsistence strategies and early human population history: An evolutionary ecological perspective. *World Archaeol.* **34**, 6–25 (2002).
- Haub, C. How many people have ever lived on earth? *Popul. Today* **23**, 4–5 (1995).
- Keinan, A. & Clark, A. G. Recent explosive human population growth has resulted in an excess of rare genetic variants. *Science* **336**, 740–743 (2012).
- Gazave, E. *et al.* Neutral genomic regions refine models of recent rapid human population growth. *Proc. Natl Acad. Sci. USA* **111**, 757–762 (2014).
- Cohen, J. E. *How Many People Can The Earth Support?* (W.W. Norton & Company, 1995).
- Kingsland, S. The refractory model: the logistic curve and the history of population ecology. *Q. Rev. Biol.* **57**, 29–52 (1982).
- Sibly, R. M. & Hone, J. Population growth rate and its determinants: an overview. *Phil. Trans. R. Soc. B* **357**, 1153–1170 (2002).
- Arim, M., Abades, S. R., Neill, P. E., Lima, M. & Marquet, P. A. Spread dynamics of invasive species. *Proc. Natl Acad. Sci. USA* **103**, 374–378 (2006).
- Kelly, R. L., Surovell, T. A., Shuman, B. N. & Smith, G. M. A continuous climatic impact on Holocene human population in the Rocky Mountains. *Proc. Natl Acad. Sci. USA* **110**, 443–447 (2013).
- Hassan, F. A. *Demographic Archaeology in Studies in Archeology* (Academic Press, 1981).
- Anderson, D. G., Maasch, K. & Sandweiss, D. H. (eds) *Climate Change and Cultural Dynamics: a Global Perspective on Mid-Holocene Transitions* (Academic Press, 2011).
- Henrich, J. Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses: the Tasmanian case. *Am. Antiq.* **69**, 197–214 (2004).
- Pinhasi, R., Fort, J. & Ammerman, A. J. Tracing the origin and spread of agriculture in Europe. *PLoS Biol.* **3**, e410 (2005).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank A. Barnosky, J. Kang, N. Rosenberg, E. Lindsey, N. Villavicencio, L. Frishkoff, K. Solari, J. Hsu, and H. Frank for conversations and feedback on the paper, as well as E. Jewett and M. Edge for discussions of statistical methods. We also greatly appreciate assistance from the Stanford Geospatial Center. This work was stimulated by discussions of an NSF-funded group (EAR 1148181) comparing the timing of megafaunal extinctions with archeological and palaeoenvironmental data in South America. We acknowledge support from NSF grant BCS 1515127, as well as NSF Graduate Research and ARCS fellowships to A.G., a Stanford Interdisciplinary Graduate Fellowship to A.M.M., and a Gabilan Fellowship to E.A.H. Radiocarbon data and associated information are available in the Supplementary Information.

Author Contributions A.G. conducted analyses and wrote the first draft of the paper. A.M.M. collected and collated the database, and conducted ArcGIS analyses. E.A.H. advised the analyses and initiated the project. All authors interpreted results, and contributed to framing and editing of the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.G. (agoldb@stanford.edu).

METHODS

Data reporting. No statistical methods were used to predetermine sample size.

Data, calibration, and quality control. Our data set of radiocarbon dates from archaeological sites in South America was gathered from previously published journal articles and online resources, including English, Spanish, French, and Portuguese publications (Supplementary Data 1). We included dates at least 2,000 ¹⁴C years old. We follow the convention of refs 12 and 30, only excluding dates that have been rejected by the field, such as outliers for early colonization (before 12,780 ¹⁴C years ago). This is expected to generate increased noise in the sample, but is likely outweighed by the increased sample size as the noise is approximately randomly distributed in the data set. We only included dates that have been associated with human activity, culture, or remains. No single sources of continent-wide databases for radiocarbon dates over this time period exist; therefore there are likely to be some omissions. Additionally, sea levels have risen since the first peopling of South America, therefore early coastal sites are not sampled.

Locations within 0.1 degrees latitude or longitude from each other are considered a single site. Sites were geo-referenced where appropriate. We excluded dates without an identifiable lab number or latitude and longitude for analyses. To buffer the effects of oversampling without substantial information loss, we considered a pruned version of the data set. Using the OxCal³¹ R_combine command, we merged radiocarbon dates that occur within 200-year intervals at site into a single normalized probability density of calibrated calendar dates^{12,15}. The SCPD for the pruned data are highly correlated ($R^2 = 0.971$, Pearson) with that of the unpruned data. Indeed, the unpruned data set provides stronger support for our two-phase model, with higher late Holocene growth rates. For all subsequent analyses and those in the main text, we use the pruned data set and assume independence of merged dates.

After these quality controls, 2,576 independent merged 'dates' at 1,147 unique sites. Radiocarbon dates and their associated errors were calibrated using OxCal 4.2 (ref. 31) and the most recent Southern Hemisphere calibration curve, SHCal13 (32).

Inclusion of shell dates. Our database includes 597 radiocarbon dates from marine shells or otoliths. Calcium carbonates from marine shells and otoliths require special calibration due to the marine reservoir effect, which results from temporal lags in carbon mixing within ocean waters. Reservoir effects can vary on local scales due to upwelling and circulation patterns, necessitating correction with both local and global corrections. We used published radiocarbon dates that were corrected for the local reservoir with a DeltaR value. These dates were then calibrated with the Marine13 Curve³³ as is standard practice. Large coastal cultures throughout the continent are known to use marine resources heavily^{4,8}; exclusion of shell dates would bias our results. Other studies of South American demography also opt to include marine shells in their analyses³⁴.

Palaeoclimate data. We used paleoclimatic data from two different sources to show general trends over large spatial scales. We consider data from ice cores sampled in Sajama, Bolivia³⁵, and Vostok, Antarctica³⁶.

Kernel density estimation. To visualize changes in site density and location through time, we created bivariate kernel density maps using the ArcGIS 10.2.2 (ref. 37) spatial analyst toolbox. We considered sites that are occupied in 1,000-year time bins, including only a single date per site within each period. Similar analyses using 500-year time bins did not change spatiotemporal patterns, and the larger archaeological sample sizes of 1,000-year time bins provide a clearer image of change through time. Thus we present the results of 1,000-year time bins here.

Kernel density estimation is widely used in spatial analyses, in which data in the form of points is used to create a density surface³⁸, based off smoothing of the surface near observed data and convolution of overlapping surfaces. We used an Epanechnikov kernel process³⁹, which more effectively minimizes mean integrated square error as compared with the alternative Gaussian kernel process^{38,40}. Magnitudes are smoothed using a bilinear interpolation.

To prevent over or under smoothing, we consider two methods to choose the search radius (bandwidth), a scaling factor that extends from each data point and determines the width of the probability mass as it spreads from that point⁴⁰: (1) independent best-fit search radius for each time slice, and (2) mean of independent best-fit radiuses applied to all time-slices.

For a given time period, with sample size n , the best-fit search radius, R , is given by³⁷

$$R = 0.9n^{-1/5} \min(D_s, \ln(2)^{-1/2} D_m),$$

where D_m is the median of distances from the mean centre of all points in the data set, and D_s is the standard distance between those points.

We first consider the best-fit radius for each time period, which reflects differences in sampling and uncertainty. The radiuses ranged from 563 km to 1145 km, with the all but the oldest two time bins falling within a 560–700 km

range, consistent with estimates for North America⁴¹. With these best-fit radiuses, we found a density range of 0.03 to 1.29 magnitude per square decimal degree.

We compared results from time-specific radiuses to a single radius for all time slices. We use the mean of all best-fit radiuses, 660 km, without outliers as a single search radius for all time periods⁴¹ (Extended Data Fig. 2). Other values for the radius considered (500, 600, 800, 1000) showed the same qualitative patterns. We used a density interval of 0.15, with nine discrete contours from 0.01–0.15 to 1.20–1.35 measured as magnitude per square decimal degree.

Summed probability distributions of calibrated dates and biases. SCPDs as a proxy for population size. The density of radiocarbon dates over time is an extensively used tool to estimate relative levels of population size^{12–15,42}. We calculated the summed probability distributions of the calibrated radiocarbon dates (SCPDs), or the sum of the distributions for each calibrated date, using the *sum* function in OxCal at the two-sigma level. That is, an SCPD calibrates each merged radiocarbon date independently, which produces a set of 2,576 probability distributions of calendar dates. These calendar dates were then treated as independent random variables and their distributions summed to give a single convolution of all dates, the SCPD.

SCPDs are beneficial because they include the uncertainty associated with calibrated dates by considering each date as a distribution rather than a point estimate. The convolution of the distributions of each calibrated date into a single SCPD requires the assumption that the dates are independent. The pruning of dates described above into 200-year bins decreases autocorrelation.

We made two additional assumptions: (1) that the amount of material dated is approximately proportional to the population size, and therefore the density of material is a proxy for population size, and (2) that the amount of material dated is approximately proportional to the amount of material initially present, that is, that sampling and taphonomic biases are not large and systematic. These assumptions can be considered reasonable because independent lines of evidence demonstrate strong correlations between SCPDs and multiple other proxies of population size over time, suggesting that SCPDs reflect demographic events. Specifically, SCPDs closely track the number of archaeological sites, density of artefacts, palaeodemographic and burial histories, and environmental-use records^{14,43–45}, though one must interpret these with care^{12–15,42,46}. While sampling biases in excavation and dating, as well as sampling of the literature to make the database, can influence the shape of the SCPD, our data set is large enough that we assume it approximates a random sampling, following recommendations in the literature^{14,42,47,48}.

SCPDs assume a constant relationship between population size and the density of radiocarbon dates. Changes in SCPD may be the result of changes in the way populations distribute or leave material. In these cases, the SCPD may not reflect population size directly, but is still providing important information about changing subsistence strategy or culture. During the time period considered, most of the continent shifted from nomadic hunter-gatherers to sedentary populations. Owing to both field sampling and our quality control filters, which bias towards uniquely occupied sites^{15,44}, hunter-gatherers are likely to leave proportionally more radiocarbon dates. For example, pre-Neolithic hunter-gatherers lived approximately 25–30 individuals to a site, with more sedentary populations in the region often reaching 200, and occasionally reaching 400 individuals^{9,27}.

Taphonomic bias. Taphonomic degradation of material as a function of time is suggested to be systematic, introducing bias. Surovell *et al.*⁴⁹ provide a correction to the SCPD for taphonomy as a function of date based on volcanic records^{14,15}. However, a singular correction is unfit for diverse climates on the scale of the South American continent^{8,15,28}. The extensive preservation of the Atacama Desert cannot be considered in the same correction as dates from the highly degrading Amazon rainforest.

Additionally, we argue that owing to the particular emphasis on sampling earlier dates in the region because it coincides with first peopling and our quality controls such as the 200-year binning, taphonomic bias is a mild problem for our data¹⁵. Taphonomic bias should also effect dates more than sites, yet we see the same trend in the density of radiocarbon dates and two measures of site occupation.

The SCPD for South America presented here does not follow the curvilinear shape typical of strong taphonomic biases, where the frequency of dates is approximately exponentially related to the date, suggesting a weaker effect of taphonomy on curve shape. For example, during the 4,000-year period of approximately constant frequency of dates (9 to 5 ka), it is unreasonable to assume that the population size was decreasing at the approximate rate of taphonomic degradation over such a long period in diverse environments across the continent.

We acknowledge that some bias will exist in our data set, particularly due to differential sampling or excavation efforts, and geographic variation in preservation. For example, modern climate and vegetation plays a role in the lack of dates from the Amazon region (both sampling efforts and degradation), but may not be indicative of the absence of people. Therefore, it is not possible to draw conclusions from the lack of data. A kernel density map of sites

(Fig. 1b) suggests the regions for which we have power. We employed a number of methods to control biases and aid interpretation for regions with coverage in our data set: (1) we pruned and merged dates that occur within 200-year bins or within 0.1 degrees latitude or longitude, reducing differences owing to funding or interest of excavation; (2) we compared date and site densities (Fig. 3), as sites are less susceptible to excavation or funding preferences; (3) we spatially smoothed date occurrences using kernel density maps (Fig. 2). Therefore, occurrence of a date in our data set is not interpreted as occupation of the site alone, but rather evidence of human presence in the local region over a 1,000-year period.

Calibration curve effects. Notably, the calibration curve itself can introduce peaks and troughs that look similar to population size changes. We took multiple steps to ensure that our results are robust to the effects of the calibration curve on SCPDs. First, we plotted a 400-year moving average of the density (Fig. 3b), as suggested by ref. 14. Additionally, we plotted histograms of the frequency of occupied sites over time, both calibrated and uncalibrated (Fig. 3c). Importantly, estimates of growth rates and population sizes are based off the best-fit curves, which do not explicitly show calibration effects. We tested 500-year and 1,000-year minimums on phase period for curve fitting, to minimize over-fitting to calibration artefacts. Analyses are focused on the calibrated date range 14,000 to 2,000 calibrated years before present, but we include dates on either side to minimize edge effects.

The South American continent crosses the equator. Therefore, our data set is comprised of dates in both the Southern and Northern Hemispheres. For our analyses, all dates are calibrated using the South Hemisphere curve³². A reasonable alternative would be to calibrate dates above the equator using IntCal13 (ref. 33). The SCPD using IntCal13 is highly correlated with our SCPD using only SHCal13 ($R=0.99$, Pearson correlation coefficient), showing no qualitative differences. This is expected, as calibration differences between the two curves are on the order of 15 years⁵⁰, and only 6.2% of dates are above the equator. The division of the two calibration curves by hemisphere is not precise. Circulatory patterns across the South American continent, and therefore ¹⁴C levels and calibration effects, are more similar to each other than northern South American patterns are to Northern Hemisphere continents, particularly as all our Northern Hemisphere dates fall within 15 degrees of the equator, where the Hadley cells function⁵¹. We therefore used a single curve for all dates, SHCal13.

Simulating constant size. As the calibration curve can introduce artificial troughs and peaks into the SCPD, we followed the method of Shennan *et al.*¹² to test if the changes in density from 9 to 5.5 ka are consistent with a constant population size or if there is evidence the population is oscillating around a constant mean. Through simulation of SCPDs, we constructed bounds for the density over time under the null hypothesis of constant size.

For a constant population size, calendar dates (not radiocarbon dates) are expected to follow a Poisson distribution. We consider the date range 9 to 5.5 ka, simulating N dates from a uniform distribution of integers between 10 and 4.5 ka to avoid edge effects. We used $n=1,121$, the number of dates in our pruned data sample with calibrated medians occurring in this time interval. Both samples were taken using the *randsample* command in MATLAB⁵². Using the *R_simulate* function in OxCal, and assigning radiocarbon errors by randomly sampling (with replacement) from errors in the database, we converted the simulated calendar dates to radiocarbon dates. Then, as with the data, we use the sum function in OxCal to create the SCPD for each simulation. Normalizing the SCPD from each simulation and from the data to 1 during the time interval from 9 to 5.5 ka, we considered the distribution of densities at 5-year time intervals. The 98% interval for 200 simulations is considered a guide rather than a distribution from which to calculate P values because we did not account for autocorrelation and multiple testing.

Quickly varying from the upper bound to lower bound of the interval for multiple cycles as observed is unlikely and may suggest oscillatory population size. Extended Data Fig. 3 shows a histogram of the variance in normalized density for each of the 200 simulation runs. The observed variance in the SCPD for South America 9 to 5.5 ka is outside the simulated distribution, and is $2.3\times$ the mean of the distribution. The variance is a conservative statistic as it does not account for clustering of densities above or below the simulated confidence interval.

A two-phase model for population size. To determine the general patterns that govern population size over time in South America, we fit 9 models of piecewise exponential and logistic growth (Extended Data Table 1) to the period between 14 to 2 ka. Evidence before ~ 14.5 ka is still controversial, so we picked this range to avoid biases and edge effects from calibration and sampling. Possible earlier colonization is not excluded based on our analyses, though lack of dense radiocarbon dates suggests any occupation would be at very low densities.

We fit the curves using *nlinfit* in MATLAB, which uses iterative least squares estimation to fit a nonlinear regression to a user-specified function. For exponential

growth, we used the two-parameter model $y=p_0e^{rx}$, where x is time, p_0 is an initial value and r is growth rate. For logistic growth, we have a three-parameter model, $y=\frac{K}{1+e^{r(x-p_0)}}$, where K is carrying capacity or the limit of the logistic function, r is the steepness of the curves, and p_0 is the midpoint of the sigmoid.

For models with multiple pieces, which we refer to as phases, we add a parameter for the change point from one curve to the next, allowing for discontinuities between neighbouring curves. To prevent over-fitting to calibration artefacts, we require phases to last at least 1,000 years, and consider change points in 500-year intervals.

The SCPD is a convolution of the distribution of calendar dates, and therefore the densities at nearby time periods are not independent. Additionally, the distributions of the calendar dates are nonstandard and vary for each calibrated date. That is, while radiocarbon dates can be represented as a normally distributed random variable, once calibrated, the probability distribution of a date is no longer an exponential family distribution, and differs for each date. Therefore we cannot simply calculate the likelihood of the data under the fit logistic and exponential models. Rather, we use a proxy for the log-likelihood to enable use of likelihood methods for parameter estimation and model choice. Specifically, we took a random sample of dates from the SCPD ($n=2,576$, the number of merged dates), and assumed sample dates are independent ($x_i \in x_1, \dots, x_{2576}$). Using the models fit to the SCPD, we treated the sampled dates, x_i , as the observed data, calculating a quantity that measures the fit of the model based on model log-likelihood calculations. That is, for exponential and logistic models, respectively, we calculated a proxy for the likelihoods as

$$L(p_0, r; x_1, \dots, x_n) = \prod_{i=1}^n p_0 e^{rx_i},$$

$$L(p_0, r, K; x_1, \dots, x_n) = \prod_{i=1}^n \frac{K}{1 + e^{-r(x_i - p_0)}}.$$

The proxy log-likelihoods and associated maximum likelihood estimates of phase-change dates for each mode are presented in Extended Data Table 1. We calculated Bayesian information criteria⁵³ to determine the best-fit model: a two-phase model with logistic growth from 14 to 5.5 ka and exponential growth from 5.5 to 2 ka. The second best model was of similar form, logistic growth followed by two phases of exponential growth, though it is not well supported by the data, with a difference in BIC from the best-fit model, ΔBIC , of 11.7.

The timing of renewed exponential growth is of demographic and cultural importance. To further examine the estimated change point for two-phase logistic-exponential model, we calculated the proxy likelihood of the model varying the date of the change from logistic to exponential growth in 100-year increments in the range 10 to 2 ka. We found similar support for dates between 5.7 to 4.1 ka ($\Delta BIC < 3$).

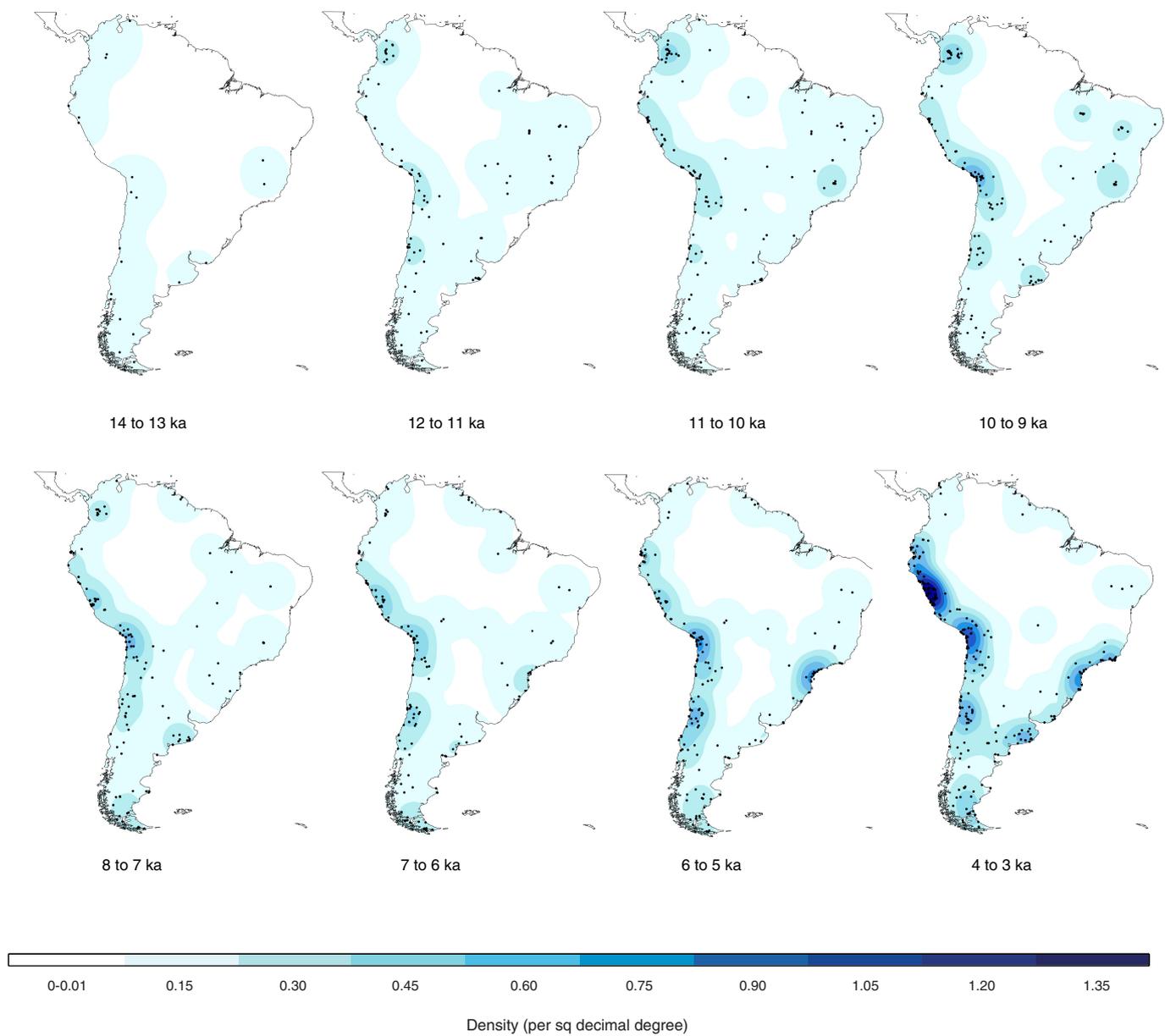
We estimate yearly growth rates over time under the model (Fig. 4a) by calculating $(y_{x+1}/y_x) - 1$. Growth rates should be interpreted as approximate, as they may be a composite of growth and taphonomic processes.

Estimating population size. Ratio of the density of radiocarbon data over time. Assuming the SCPD is proportional to population size over time, the ratio of ending to starting density is a measure of the relative increase in population size over that time period. To buffer edge effects, especially as the earliest rare dates are a poor measure of population size, we considered instead the mean density of the 500 years surrounding 13.5 ka to the 500-year mean surrounding 2.5 ka. That is, we calculate the ratio of the mean density from 13.75 to 13.25 ka to the mean density from 2.75 to 2.25 ka. This ratio is 993; interpreted as the population size at 2.5 ka was 993 times larger than that at 13.5 ka.

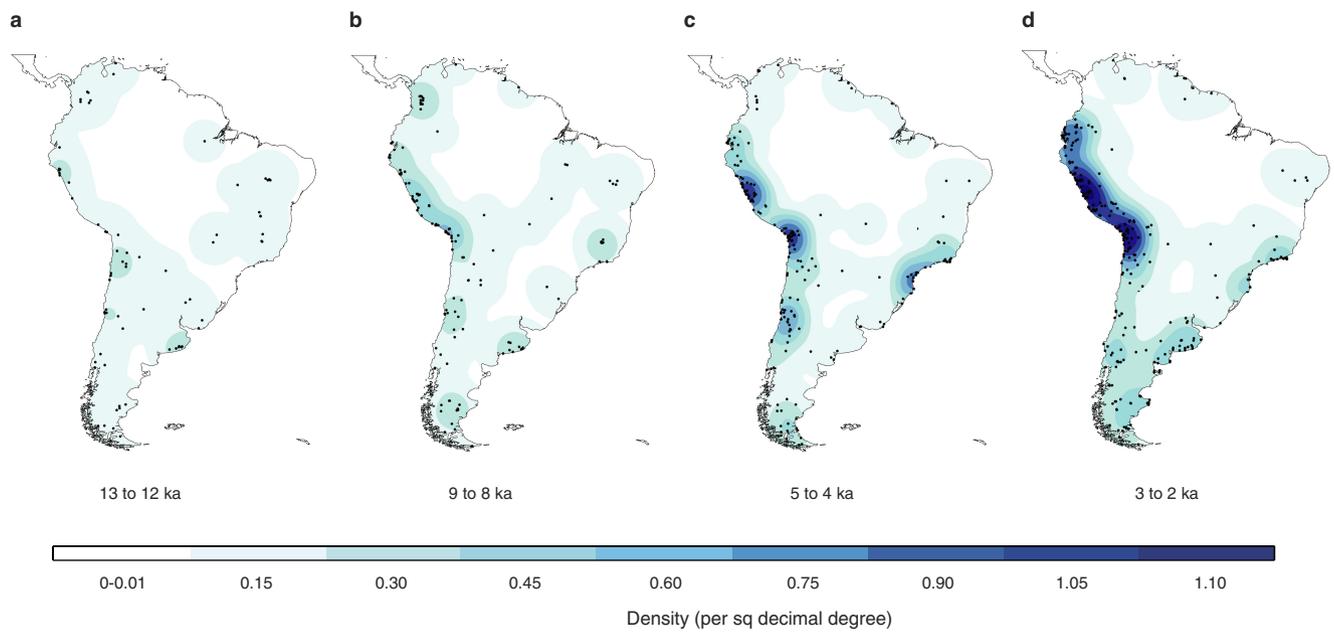
Population size under the two-phase model. As the number of individuals initially entering South America is uncertain, we considered the N -fold growth over time, rather than the census size. We defined N_x as the population size at time x , and have N_x/N_0 as the ratio of population size at time x to initial population size, estimated under the model.

Under the two-phase model, we use the calculated growth rates to iteratively estimate population size over time. We fit the model starting at 14 ka to avoid edge effects of the SCPD, but securely dated archaeological sites place colonization of South America at ~ 14.8 ka or before¹⁻⁵. Therefore, we extended the fit logistic model to 14.8 ka, starting with a population size $N_0=1$ individual, and multiplying the current size by the growth rate at that time. That is, the population size at time x is given by $N_x = N_{x-1}(r_x + 1)$, for growth rate r , which is equal to $N_x = N_{x-1}(y_x/y_{x-1})$. Carrying capacity for South America was reached approximately 8.5 to 9 ka, with $N_x/N_0=274$. By 2 ka, $N_x/N_0=616$.

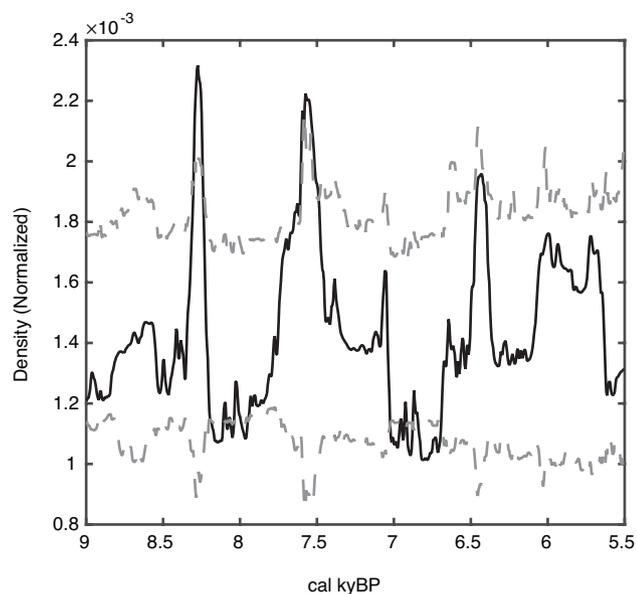
31. Ramsey, C. B. Bayesian analysis of radiocarbon dates. *Radiocarbon* **51**, 337–360 (2009).
32. Hogg, A. G. *et al.* SHCal13 Southern Hemisphere calibration, 0–50,000 years cal bp. *Radiocarbon* **55**, 1889–1903 (2013).
33. Reimer, P. J. *et al.* IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal bp. *Radiocarbon* **55**, 1869–1887 (2013).
34. Gayo, E. M., Latorre, C. & Santoro, C. M. Timing of occupation and regional settlement patterns revealed by time-series analyses of an archaeological radiocarbon database for the South-Central Andes (16°–25° S). *Quat. Int.* **356**, 4–14 (2015).
35. Thompson, L. G. *et al.* A 25,000 year tropical climate history from Bolivian ice cores. *Science* **282**, 1858–1864 (1998).
36. Petit, J. R. *et al.* Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* **399**, 429–436 (1999).
37. ArcGIS Version 10 (Redlands CA: Environmental Systems Research Institute Inc., 2010).
38. Silverman, B. W. *Density Estimation for Statistics and Data Analysis* (Chapman & Hall, London, 1986).
39. Epanechnikov, V. A. Non-parametric estimation of a multivariate probability density. *Theory Probab. Appl.* **14**, 153–158 (1969).
40. Wand, M. P. & Jones, M. C. *Kernel Smoothing* (Crc Press, 1994).
41. Chaput, M. A. *et al.* Spatiotemporal distribution of Holocene populations in North America. *Proc. Natl Acad. Sci. USA* **112**, 12127–12132 (2015).
42. Brown, W. A. Through a filter, darkly: population size estimation, systematic error, and random error in radiocarbon-supported demographic temporal frequency analysis. *J. Archaeol. Sci.* **53**, 133–147 (2015).
43. Hinz, M., Feeser, I., Sjögren, K. G. & Müller, J. Demography and the intensity of cultural activities: an evaluation of Funnel Beaker Societies (4200–2800 cal bc). *J. Archaeol. Sci.* **39**, 3331–3340 (2012).
44. Downey, S. S., Bocaage, E., Keri, T. E., Edinborough, K. & Shennan, S. The Neolithic demographic transition in Europe: correlation with juvenility index supports interpretation of the summed calibrated radiocarbon date probability distribution (SCDPD) as a valid demographic proxy. *PLoS ONE* **9**, e105730 (2014).
45. Woodbridge, J. *et al.* The impact of the Neolithic agricultural transition in Britain: a comparison of pollen-based land-cover and archaeological ¹⁴C date-inferred population change. *J. Archaeol. Sci.* **51**, 216–224 (2014).
46. Contreras, D. A. & Meadows, J. Summed radiocarbon calibrations as a population proxy: a critical evaluation using a realistic simulation approach. *J. Archaeol. Sci.* **52**, 591–608 (2014).
47. Michczynska, D. & Pazdur, A. Shape analysis of cumulative probability density function of radiocarbon dates set in the study of climate change in the Late Glacial and Holocene. *Radiocarbon* **46**, 733–744 (2004).
48. Michczynska, D. J., Michczynska, A. & Pazdur, A. Frequency distribution of radiocarbon dates as a tool for reconstructing environmental changes. *Radiocarbon* **49**, 799–806 (2007).
49. Surovell, T. A., Finley, J. B., Smith, G. M., Brantingham, P. J. & Kelly, R. Correcting temporal frequency distributions for taphonomic bias. *J. Archaeol. Sci.* **36**, 1715–1724 (2009).
50. Reimer, P. J. *et al.* IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal bp. *Radiocarbon* **55**, 1869–1887 (2013).
51. Ogburn D. E. Reconceiving the chronology of Inca imperial expansion. *Radiocarbon* **54**, 219–237 (2012).
52. Seidel, D. J., Fu, Q., Randel, W. J. & Reichler, T. J. Widening of the tropical belt in a changing climate. *Nature Geosci.* **1**, 21–24 (2008).
53. MATLAB Release 2015a (The MathWorks, Natick, Massachusetts, USA).
54. Schwarz, G. E. Estimating the dimension of a model. *Ann. Stat.* **6**, 461–464 (1978).



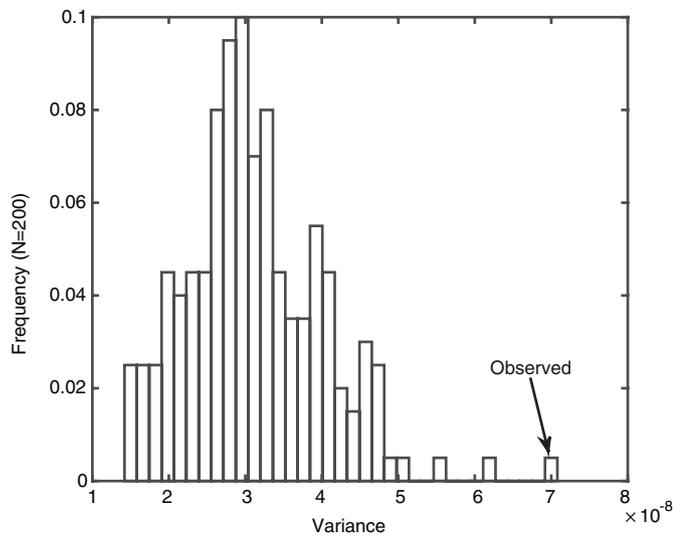
Extended Data Figure 1 | Kernel Density estimates of occupied area for alternative time bins. Estimates of occupied area in 1,000-year time bins for dates not present in Fig. 2, considering uniquely occupied sites in each time bin.



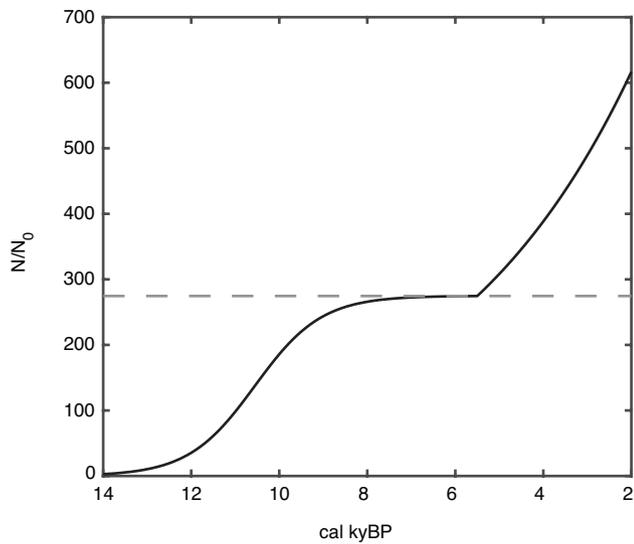
Extended Data Figure 2 | Kernel Density estimates of occupied area for fixed search radius. Estimates of occupied area in 1,000-year time bins for dates present in Fig. 2, using a fixed radius for all time slices of 660 km.



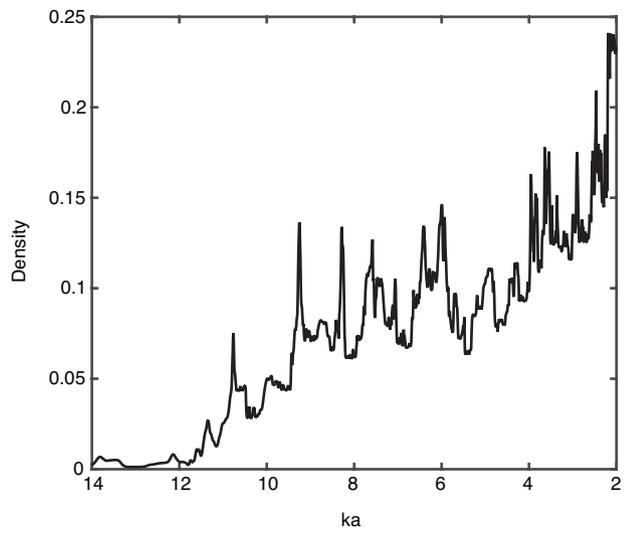
Extended Data Figure 3 | Statistical evaluation of constant population size in the mid-Holocene. The simulated 98% confidence interval (grey) for radiocarbon density under a constant population size in 5-year bins, with the observed SCPD for the data (black).



Extended Data Figure 4 | Observed oscillations are outside the variation of simulated constant size. A histogram of the variance of 200 simulated constant size SCPDs, with the observed variance of density in the observed SCPD from 9 to 5.5 ka outside the distribution of those simulated.



Extended Data Figure 5 | Inferred population size over time. Assuming colonization of South America 14.8 ka by a population of size N_0 , we plot the relative increase in population size over time under the two-phase model (black). Carrying capacity (grey) occurs approximately 8.5 ka with a relative population size $N/N_0 = 274.6$, and a final population size at 2 ka of $N/N_0 = 616.6$.



Extended Data Figure 6 | SCPD for Andean/Pacific Coastal Region.
The Pacific Coast region and Andes, which represents 52% of dates and multiple of the major cultural centres, show a similar trend as the SCPD for the continent.

Extended Data Table 1 | Model likelihoods and model choice

Model	Transitions (ka)	Log likelihood	BIC	Δ BIC
Single phase, exponential	N/A	-4335.45	8686.56	264.91
2 phases, exponential	10	-4208.63	8448.59	26.93
3 phases, exponential	10; 5.5	-4195.09	8437.18	15.52
4 phases, exponential	11; 7; 5.5	-4190.90	8444.46	22.81
5 phases, exponential	11; 7; 6.5; 5	-4186.96	8452.24	30.59
6 phases, exponential	11; 7; 6.5; 5.5; 4	-4185.06	8464.12	42.47
Single phase, logistic	N/A	-4241.69	8506.88	85.23
2 phases, logistic then exponential	5.5	-4191.25	8421.65	0
3 phases, logistic then 2 exponential	4; 3	-4189.27	8433.36	11.71

Nine piece-wise exponential and logistic models considered, with associated log-likelihoods and Bayesian information criteria. Best-fit model is emphasized: logistic growth until 5.5 ka followed by exponential growth until 2 ka.