Supporting Information

Kohler and Reese 10.1073/pnas.1404367111

SI Text

Sources of Data

This expansion to ref. 1 extends the spatial window within which $_{15}P_5$ values were collected or calculated to encompass the entirety of the pre-Hispanic Southwest. We also assembled life table data within this window where we could find it. We attempted to locate assemblages collectively spanning as much of the total occupation span for the region as possible. Although we looked for assemblages predating ~1100 B.C., we could not generate an adequate sample, which argues for high mobility and low population sizes before that date. Mortuary assemblage data were collected from as many sources as could be located, including published articles, journals, volumes, published and unpublished site reports, and databases provided by colleagues. We accepted all data from these sources with the following exceptions:

- We excluded assemblages where we could not estimate their dates or age distributions to our satisfaction;
- We excluded assemblages obviously affected by massacres or extreme perimortem processing possibly indicating cannibalism;
- Unless they could be used as part of a composite assemblage, we generally excluded isolated burials;
- We excluded one $_{15}P_5$ value of 0.954 from a large assemblage (Snaketown, A.D. 1000; Table S1) as an extreme positive outlier (Fig. S1).

In the northern Southwest, the dates we assign to assemblages are typically derived directly (or indirectly, via ceramics that have previously been calibrated) from tree-ring chronologies. In the desert Southwest, most dates are directly or indirectly from ^{14}C determinations. There is no standard repository for archaeological or bioarchaeological data so it is impossible to census the population of excavated pre-Hispanic skeletal materials or to estimate its size. Most of these materials have been repatriated and are unavailable for new study. This study increases the number of sites/composite assemblages by 380% and the number of human remains by 230% relative to ref. 1.

Coding Conventions

Tables S1 and S2 contain many entries for skeletal counts that are not integers. Many sources offered age estimates cross-cutting the age categories we used (5–19 and ≥20). In those cases, their counts were distributed proportionately among our age ranges. For example, an individual aged 18–21 would be distributed as 0.5 individuals in the 5–19 category and 0.5 individuals in the ≥20 category. We attempted to accurately interpret an individual's actual age when researchers used nonquantitative characterizations such as infant, adolescent, etc., using other information specific to each source. All analyses include both aged cremations and aged inhumations. Cremations are most common in the Sonoran Desert and Tonto Basin sequences, between ~A.D. 700 and 1300.

Statistical and Graphic Conventions

Although the original presentation of the Neolithic Demographic Transition for Europe and the Levant assigned assemblages a chronological placement relative to the local assumption of a Neolithic way of life (2), here we use absolute dates. This is possible because of the smaller area examined, and results in a clearer presentation.

In Fig. 1, site/composite sample symbols are proportional in size to number of individuals aged ≥5. Regional affiliations of sites were assigned using information in each source. Composite data points were only generated from assemblages within a single region, and close together in time. Their coordinates were found by locating each contributing site and assigning the composite an average location, weighted according to the sample sizes of each contributing site.

Loess analyses (Figs. 2–4) were weighted by number of individuals aged ≥5 in each assemblage, making aggregation or deletion of small assemblages less critical than for an analysis not so weighted. The date assigned to each assemblage was the midpoint assigned to that assemblage in each source, or our best estimate of what that midpoint would be when the source was not explicit. The analyses in Fig. 2 used all samples with ≥5 individuals aged ≥5 y and a smoothing parameter α of 0.45. The analyses in Fig. 3 and the analysis generating e_0BA in Fig. 4 used all samples with ≥10 individuals and a smoothing parameter of 0.35. All loess fits used degree = 1 (linear) for the local polynomials in each local regression, and cubic interpolation for blending local polynomial fits at the kd tree vertices. Confidence intervals in Figs. 2–4 are 90% around the fit. These intervals integrate the effects of uncertainty due to variable and often small sample sizes and sample variability in values for the dependent variable that are close together in time.

In Figs. 2–4, we place vertical references lines at A.D. 1 and 1280 to mark the beginning of the first millennium A.D. and the time by which the depopulation of the northern Southwest was essentially complete. In Fig. 2, the horizontal reference lines at $_{15}P_5 = 0.18$ mark Bocquet-Appel's estimate for the value of $_{15}P_5$ in a stationary population. This estimate is probably too high for the Southwest, because the Sonoran Desert populations were obviously surviving, and probably growing slowly, despite the fact that that their $_{15}P_5$ values did not exceed 0.18 on average until nearly A.D. 1.

Flows of populations between subregions, likely large in the A.D. 1200s, might subtly affect the juvenility indices in Fig. 2 by depression in the source area and inflation in the sink. However, the Kayenta region is a probable source for migration to the Tonto Basin and Sonoran Desert, and yet those destinations evidence very low juvenility indices in the 1200s and early 1300s. In any case, these effects, if they exist, should largely disappear when the Southwest is considered in its entirety (Fig. 3), assuming that immigration into and out of the entire region was minimal.

Cremation Analysis

To assess whether the proportion of cremations in some of these assemblages affects their $_{15}P_5$ values, we regressed $_{15}P_5$ on proportion cremations [n cremations/(n cremations + n inhumations)] where both quantities include only individuals ≥5 y old, using the 15 sites with at least 1 aged cremation identified in Table S1. We used weighted least squares with weights proportional to sample size (total aged burials ≥5 y old, both cremations and inhumations). There is no significant linear relationship between the two quantities ($_{15}P_5 = 0.288–0.123$ * proportion cremations; $r^2 = 0.07$; $P > F = 0.33$).

Crude Birth Rates

For compatibility with earlier research (2), we define the crude birth rate (CBR) as the number of living births per year divided by the current population size (births per person per year). We estimate birth rate from $_{15}P_5$ according to the formula provided in ref. 2, table 2:

$$\mathrm{CBR} = 0.00375 + 0.15334({}_{15}\mathrm{P}_5)^{0.89074}.$$

CBR is more often presented as births per 1,000 people per year. The CBR we estimate at the beginning of our sequence (0.02) is equivalent to 20 births per 1,000, slightly above the estimated world rate of 18.9 births per 1,000 in 2013 (3). Our estimated Southwest-wide CBR for the A.D. 500–1100 period exceeds, although its confidence interval includes, CBR estimates for the fastest-growing countries in the world in 2013, Niger (46.8) and Zambia (42.8) (3).

Life Expectancy Calculations

In addition to the relatively random processes affecting recovery of all materials from the archaeological record, life expectancy calculations based on death assemblages are beset by systematic biases. They depend on life tables that assume stable populations, a single population, equal probability of enumeration in all age categories (hence, good preservation of younger individuals) and accurate aging. All of these assumptions are potentially problematic here and for most other skeletal populations (4, 5). Proposed remedies for older-adult age underestimation (6) require reanalysis of collections, an impossibility here. In most cases, e_{15} values presented here (e_{15}LT where LT indicates they were calculated from life tables) were obtained directly from the literature. For four sites noted in Table S1, we computed e_0LT and e_{15}LT from age distributions given in the reports using the R program ArchLT(), provided by its author, Tim Riffe (Department of Demography, University of California, Berkeley). For San Cristóbal, we used values of e_{15} computed by this program using ages recently estimated from Native American Graves Protection and Repatriation Act forms from the American Museum of Natural History rather than those available for this site in (7). Fig. 4 and Tables S1 and S3 add 15 y to the e_{15} estimates to make them comparable to the average age at death estimate from e_0.

Because of these problems, we also present a calculation of life expectancy at birth (e_0LT), where LT indicates that it was calculated in the standard way, from life tables, and add an alternative estimate for life expectancy at birth, e_0BA, based not on life tables but computed from the juvenility index using the relationship between that index, e_0, and r computed from 45 reference life tables from various world locations in the 17th to 20th centuries (2). The advantages of this approach are that it avoids the infant underestimation and adult-age underestimation affecting life table approaches. The disadvantages are that we must assume that mortality patterns in these relatively recent populations are similar to those in the US Southwest in much more ancient times, and that we must hypothesize a value for r. We assumed r = 0.005 before A.D. 1300, and r = 0 after that. Although these two estimates for e_0 often agree in direction of trend, e_0BA is more variable through time.

1. Kohler TA, Glaude MP, Bocquet-Appel J-P, Kemp BM (2008) The Neolithic Demographic Transition in the U.S. Southwest. *Am Antiq* 73(4):645–669.
2. Bocquet-Appel J-P (2002) Paleoanthropological traces of a Neolithic Demographic Transition. *Curr Anthropol* 43(4):637–650.
3. CIA (2014) *Field Listing: Birth Rate*. Available at https://www.cia.gov/library/publications/the-world-factbook/rankorder/2054rank.html. Accessed June 20, 2014.
4. Alesan A, Malgosa A, Simó C (1999) Looking into the demography of an Iron Age population in the western Mediterranean. I. Mortality. *Am J Phys Anthropol* 110(3):285–301.
5. Milner GR, Wood JW, Boldsen JL (2008) Advances in paleodemography. *Biological Anthropology of the Human Skeleton*, eds Katzenberg MA, Saunders SR (Wiley, New York), 2nd Ed, pp 561–600.
6. Milner GR, Boldsen JL (2012) Transition analysis: A validation study with known-age modern American skeletons. *Am J Phys Anthropol* 148(1):98–110.
7. Stodder ALW (1990) *Paleoepidemiology of Eastern and Western Pueblo Communities in Protohistoric New Mexico*. PhD dissertation (Univ of Colorado, Boulder).

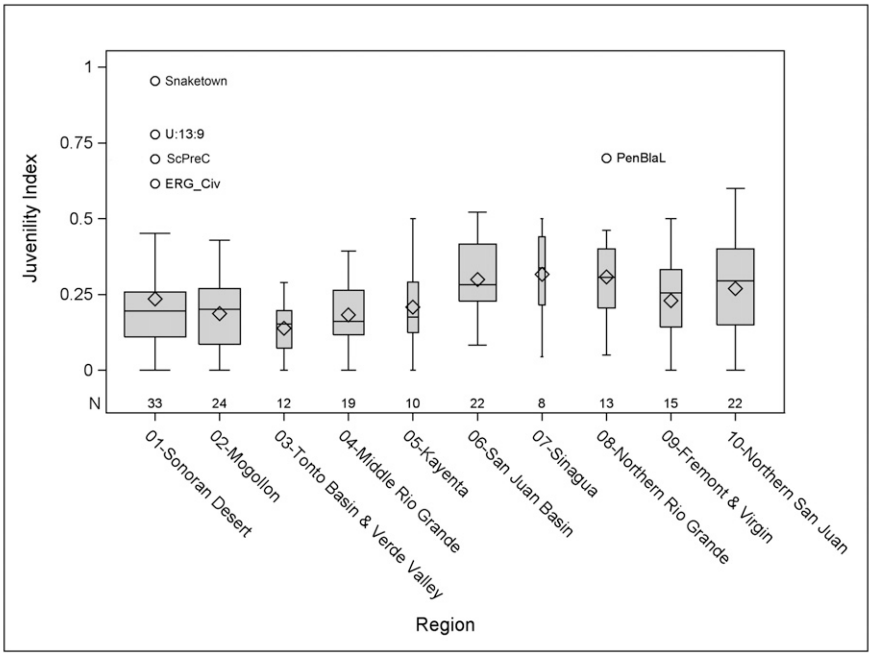

Fig. S1. Box plot, ${}_{15}P_5$ values by region. Width of boxes is proportional to number of assemblages in each region. Named outliers are more than 1.5 midspreads above the upper hinge. Snaketown is the most extreme outlier in the dataset and was eliminated from other calculations in this paper.

Other Supporting Information Files

Table S1 (DOCX)
Table S2 (DOCX)
Table S3 (DOCX)