

# Minimum founding populations for the first peopling of Sahul

Corey J. A. Bradshaw<sup>1,2\*</sup>, Sean Ulm<sup>1,2,3</sup>, Alan N. Williams<sup>2,4,5</sup>, Michael I. Bird<sup>1,2,6</sup>,  
Richard G. Roberts<sup>1,2,7</sup>, Zenobia Jacobs<sup>2,7</sup>, Fiona Laviano<sup>1</sup>, Laura S. Weyrich<sup>2,8</sup>, Tobias Friedrich<sup>1,9</sup>,  
Kasih Norman<sup>2,7</sup> and Frédéric Saltré<sup>1,2</sup>

**The timing, context and nature of the first people to enter Sahul is still poorly understood owing to a fragmented archaeological record. However, quantifying the plausible demographic context of this founding population is essential to determine how and why the initial peopling of Sahul occurred. We developed a stochastic, age-structured model using demographic rates from hunter-gatherer societies, and relative carrying capacity hindcasted with LOVECLIM's net primary productivity for northern Sahul. We projected these populations to determine the resilience and minimum sizes required to avoid extinction. A census founding population of between 1,300 and 1,550 individuals was necessary to maintain a quasi-extinction threshold of  $\lesssim 0.1$ . This minimum founding population could have arrived at a single point in time, or through multiple voyages of  $\geq 130$  people over  $\sim 700$ – $900$  years. This result shows that substantial population amalgamation in Sunda and Wallacea in Marine Isotope Stages 3–4 provided the conditions for the successful, large-scale and probably planned peopling of Sahul.**

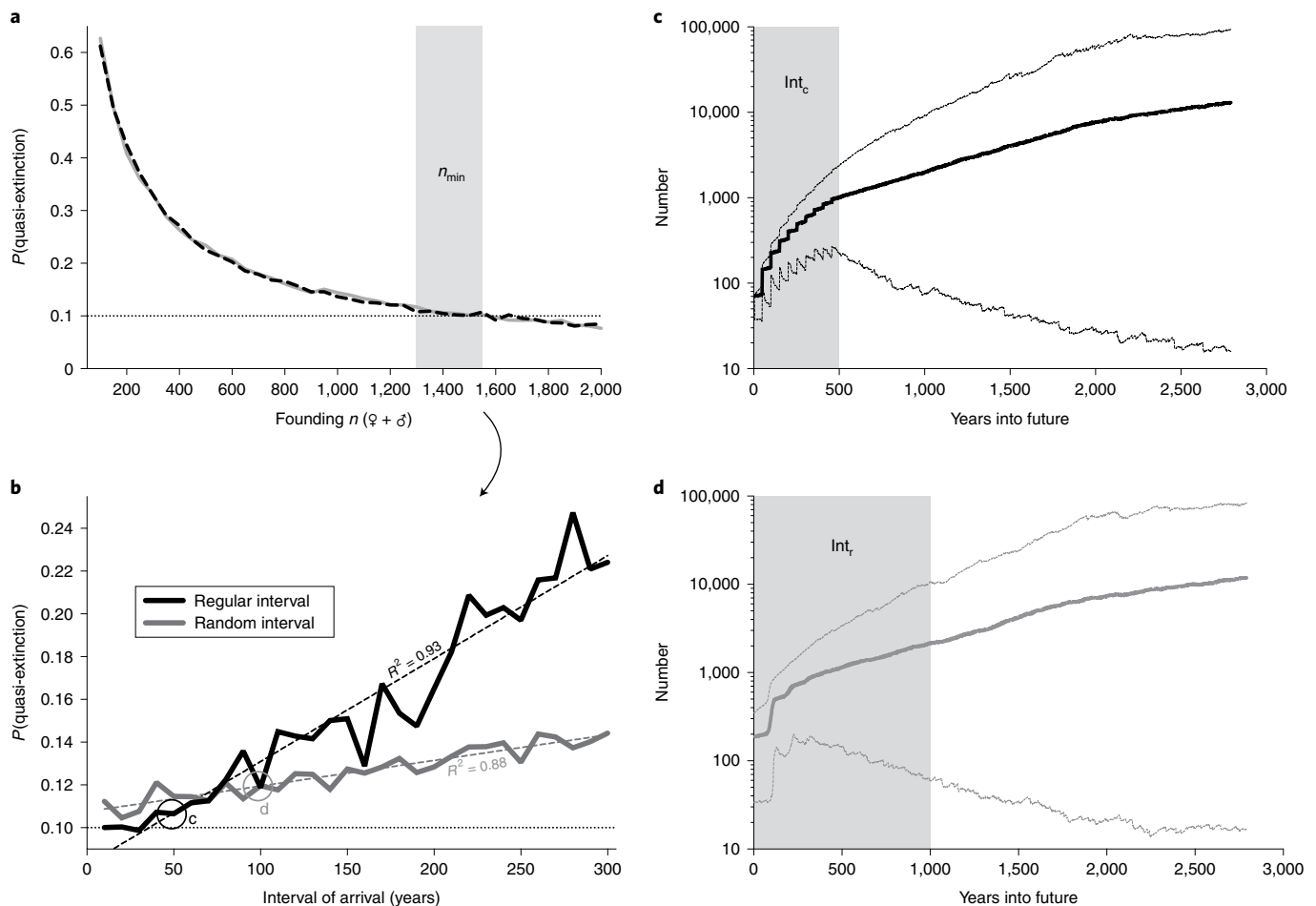
An understanding of the demographic circumstances and ecological repercussions of the arrival of the first people to Sahul (mainland Australia, Tasmania and New Guinea joined at times of lower sea level) in the Late Pleistocene remains elusive<sup>1–3</sup>. Some have previously proposed that the first peopling arose from only small family groups consisting of  $< 150$  people<sup>4–7</sup>, while more recent human behavioural ecology models suggest that several hundred people would have been required for long-term survival<sup>7</sup>. These estimates are largely speculative, but genomic research and radiocarbon-inferred demographic modelling support the larger values, and imply that population sizes of 1,000–3,000 people were more likely<sup>8–10</sup>. Quantifying the plausible demography of the first humans (that is, the rate of population changes relative to regional carrying capacity, and the duration over which populations could have persisted at low density following first arrival) is essential to ascertain the extent to which increasing human populations could have altered their environments. Quantifying demographic transitions can also potentially help to interpret the likelihood of discovering archaeological evidence, given the persistence of small human populations over extended windows of time so long ago<sup>6</sup>.

The oldest archaeological evidence claimed in Australia dates to  $65.0 \pm 5.7$  thousand years ago (ka) (95% confidence interval (CI)) for the Madjedbebe rockshelter in Arnhem Land<sup>11,12</sup>, and an increasing number of early sites have been reported dating to around or before 47 ka<sup>13–21</sup>. We therefore take the broad interval of 65–50 ka as the probable arrival window of people into Sahul. Arrival probably occurred somewhere in the northern regions of Sahul, closest to the islands of Wallacea in the Timor Sea, the Arafura Sea, and across the Torres Strait and Carpentarian Plain during times of lower sea levels<sup>22–27</sup>.

However, given that only a few archaeological sites from the north have revealed cultural remains within the putative arrival window, and that ancient DNA is poorly preserved in this region of the world, it is difficult to estimate when and where enough people first arrived in Australia to produce one of the longest standing, successful human populations in the world outside Africa—a population that went on to adapt successfully and populate the entire continent over the following several thousand years<sup>10,28,29</sup>. Recent studies have modelled plausible routes, and therefore the potential geographic locations that would have supported a successful migration to Australia<sup>22–27</sup>, but the numbers and diversity of humans who first arrived on the continent remain largely unknown<sup>10,28,29</sup>. Fundamentally then, the size and migration patterns of founding populations (that is, whether arrival was accidental by a small band of hunter-gatherers, or something larger scale and more complex) directly contributes to our understanding of modern human societies at this time, and how the peopling of Sahul fits into this broader story.

To determine the probable range of these unknown demographic conditions, we developed a stochastic, age-structured demographic model for ancient Australians to: (1) estimate the minimum size of a founding population that would be required to avoid a high risk of extinction at the time of first arrival; and (2) calculate the interval and frequency of smaller introductions that would maintain a low probability of extinction over the initial arrival window (65–50 ka). Our model is based on realistic estimates and assumptions of hunter-gatherer demography, as well as a reconstruction of the carrying capacity based on hindcasted estimates of net primary production. We hypothesize that several thousand individuals arriving over a defined period within several centuries were required to

<sup>1</sup>Global Ecology, College of Science and Engineering, Flinders University, Adelaide, South Australia, Australia. <sup>2</sup>ARC Centre of Excellence for Australian Biodiversity and Heritage, Wollongong, New South Wales, Australia. <sup>3</sup>College of Arts, Society and Education, James Cook University, Cairns, Queensland, Australia. <sup>4</sup>Climate Change Research Centre, School of Biological, Earth and Environmental Sciences, The University of New South Wales, Sydney, New South Wales, Australia. <sup>5</sup>Extent Heritage Pty Ltd, Sydney, New South Wales, Australia. <sup>6</sup>College of Science and Engineering, James Cook University, Cairns, Queensland, Australia. <sup>7</sup>Centre for Archaeological Science, School of Earth, Atmospheric and Life Sciences, University of Wollongong, Wollongong, New South Wales, Australia. <sup>8</sup>Australian Centre for Ancient DNA, University of Adelaide, Adelaide, South Australia, Australia. <sup>9</sup>Department of Oceanography, School of Ocean and Earth Science and Technology, University of Hawai'i at Mānoa, Honolulu, HI, USA. \*e-mail: [corey.bradshaw@flinders.edu.au](mailto:corey.bradshaw@flinders.edu.au)



**Fig. 1 | Estimating quasi-extinction probability for the first Sahul population.** **a**, Probability of quasi-extinction (<25 females or <50 total individuals), expressed as a function of the size of a one-off founding population ( $n$ ) according to 10,000 runs of the stochastic demographic model. The dashed black line indicates the probability decay curve assuming that the year of initial arrival fell between 65 and 55 ka. The grey line is the curve derived from an initial arrival window of 60–50 ka. The shaded  $n_{\min}$  area indicates the range of minimum founding population sizes giving  $P(\text{quasi-extinction}) = -0.1$ , which we applied in the simulations shown in **b** (symbolized by the downward-pointing curved arrow from **a** to **b**). **b**,  $P(\text{quasi-extinction})$  as a function of an increasing interval between regularly spaced arrival events (10, 20, 30, ... 300 years), each comprising one-tenth of the total founding population of 650–775 females (black line, regular interval), or as a function of randomly sampled introduced-population sizes and randomly sampled intervals averaging 10, 20, 30, ... 300 years (grey line, random interval). Also shown are the least-squares linear-regression  $R^2$  coefficients for both trajectories. The circles indicate example projections, as shown in **c** and **d**. **c**, An example 50-year constant-interval simulation occurring over 500 years ( $\text{int}_c$ ). The upper and lower lines indicate the 95% CIs of the median (darker middle line) projected population size ( $n$ ). **d**, An example 100-year random-interval simulation occurring over an average of 1,000 years ( $\text{int}_r$ ). The lines are as described in **c**.

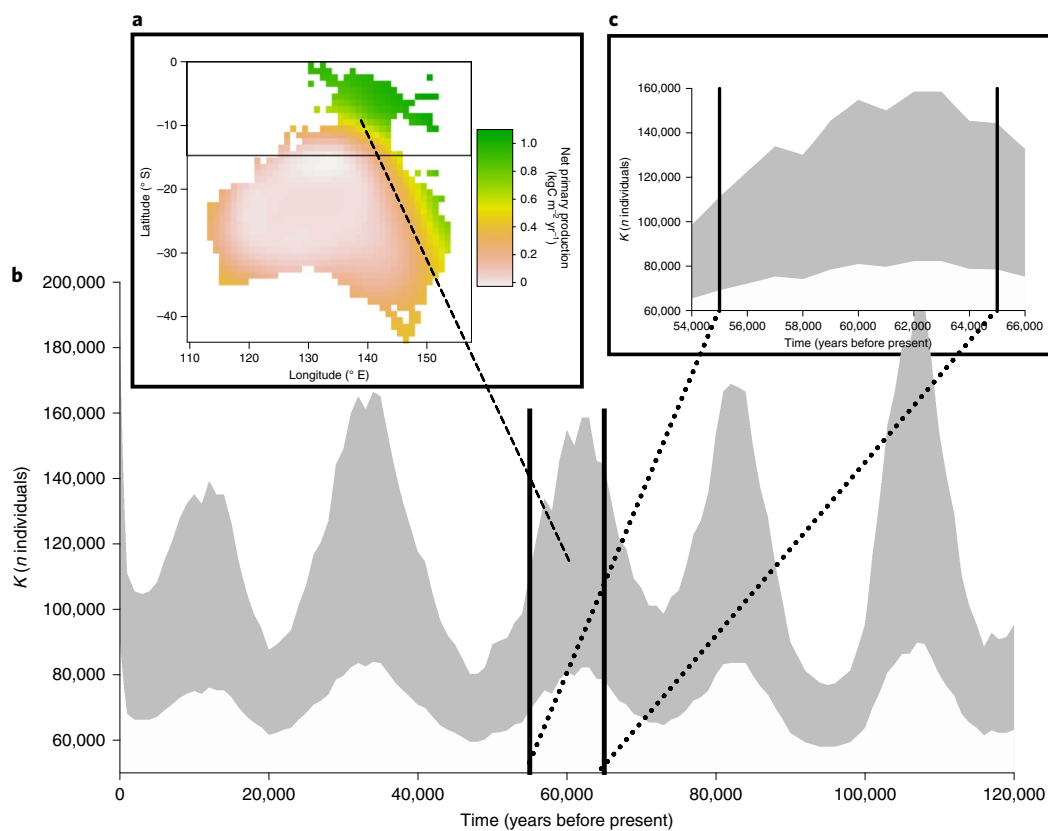
avoid extinction within the first 100 human generations following first arrival.

## Results

**Deterministic matrix properties.** The base matrix  $M$  using the Siler hazard model<sup>30</sup> to estimate the survival vector (Supplementary Fig. 1) produced a dominant eigenvalue  $\lambda = 1.0037$ , which equates to an instantaneous rate of population change ( $r$ ) = 0.0037. Applying different underlying parameters for the Siler hazard model for ‘average forager-horticulturist’ and ‘Northern Territory Aborigines’<sup>31</sup> increased the base matrix’s dominant eigenvalue ( $\lambda = 1.0085$  and 1.0201, respectively). However, given the assumed hunter-gatherer mode of subsistence during the time of initial arrival into Sahul, and the rapid rate of increase in the Aboriginal population data (collected 1958–1960) that probably under-reported infant deaths<sup>31</sup>, we maintained the lowest  $\lambda$  for the ‘average hunter-gatherer’ scenario in all subsequent simulations. This model gives a ratio of the number of female offspring in year

$t + 1$  to the number born in the previous year ( $R_0$ ) of 1.11, and a mean generation time ( $G$ ) of 27.7 years, which agrees well with the ~29-year generation length estimated from genealogy-based studies of hunter-gatherers<sup>32</sup>. Life expectancy ( $e_x$ ) according to this model increases from around 31 to a maximum 42 years between the ages of 1 and 5 years old, after which it declines linearly with age (Supplementary Fig. 1). Thus, a 20-year-old has  $e_x = 35$  (additional) years of expected life, and a 40-year-old has  $e_x = 24$  additional years of expected life (Supplementary Fig. 1).

**Minimum founding population size.** The probability of quasi-extinction ( $n < 50$  individuals or <25 females) stabilizes at around 0.1 over 100 generations for founding population sizes of 1,300–1,550 individuals (Fig. 1), or between 650 and 775 females assuming an equal sex ratio. There is quantitatively no difference between the two curves assuming different timings of first arrival (that is, 65–55 ka or 60–50 ka; Fig. 1a). The relationship between the probability of quasi-extinction and founding population size takes into consideration all uncertainty



**Fig. 2 | Change in net primary production and indicative human carrying capacity. a**, Net primary production hindcasted by the LOVECLIM<sup>50</sup> Earth system model<sup>51</sup>, showing an example for Sahul at 60,000 ka. The outlined box at the top in **a** indicates the  $1^{\circ} \times 1^{\circ}$  grid cells ( $n=166$ ) covering ‘northern’ Sahul (latitude:  $0^{\circ}$ – $14^{\circ}$  S) used to derive the relative human carrying capacity ( $K$ ) used in subsequent analyses (see Methods). **b**, The shaded area indicates the range between the 25th and 75th percentiles of  $K$  from 120 ka to the present, expressed in terms of total human population size ( $n$ ). **c**, Same as in **b**, but focusing on the period of first arrival, 65–55 ka (as indicated by vertical lines in **b**).

associated with the hindcasted carrying capacity  $K$  (Fig. 2), start year and error (process and sampling) in the Leslie matrix’s demographic-rate elements (survival and fertility; Supplementary Figs. 1 and 2), but it assumes a single-year introduction event (that is, all founding individuals arrive during the same year).

Taking this range (650–775 founding females) and dividing it by 10, such that one-tenth of this minimum arrive at incrementing intervals from 10–200 years, the probability of extinction rises approximately linearly with increasing interval length (Fig. 1b), but remains near 0.1 for up to about 70-year intervals (that is, 65–75 founding females arriving every 70 years over 7 centuries) (Fig. 1b). Assuming a non-regular (random) arrival frequency and a Gaussian-resampled arriving population size, the rate of increase in  $P(\text{quasi-extinction})$  is also linear, but less than the rate based on a regular arrival frequency (Fig. 1b); for the latter, however, the probability remains  $\sim 0.1$  up to approximately 70- to 90-year arrival intervals on average.

**Global sensitivity analysis.** The boosted-regression tree emulator for the 12-dimension, Latin hypercube-sampled parameter space indicated that the dominant (negative) influence on the probability of quasi-extinction was variation in age-independent mortality ( $b_1$  in the Siler hazard model) (Supplementary Fig. 3). Total fertility ( $F$ ) had the next-highest relative (negative) influence, followed by the (positive) influence of infant mortality rate ( $a_1$ ), and the (positive) influence of the rate of mortality decline ( $a_2$ ), with all other parameters considered having relatively weak or no detectable influence on  $P(\text{quasi-extinction})$  (Supplementary Fig. 3).

## Discussion

Discerning the plausible demographic conditions of human arrival to Australia is problematic because of the deep age of the event(s), differential preservation of archaeological material since that time, limited ancient DNA evidence, uncertainties associated with dating, taphonomic biases, and incomplete temporal and spatial coverage of samples. However, stochastic demographic models built from realistic human demographic rates, hindcasts of indicative regional carrying capacity, and relevant archaeological and genetic data to guide inference, return ecologically credible conditions. Future archaeological and palaeoecological data could also potentially exclude ecologically unrealistic hypotheses. Using the best data available (while acknowledging that having more is desirable), our models estimate that as few as  $\sim 650$  females (representing  $\sim 1,300$  individuals in total) arriving in small groups ( $\sim 130$  each) over as many as 700–900 years would be sufficient to avoid a high probability of extinction during the probable environmental conditions that dominated northern Sahul between 65 and 50 ka.

These numbers of people—whether  $\sim 130$  arriving at semiregular intervals<sup>10</sup> over a long period or 1,300 at one time—are substantively larger than ethnographically observed Aboriginal hunter-gatherers, except during larger ceremonial gatherings<sup>33</sup>. This therefore suggests that large populations were probably present in Wallacea during Marine Isotope Stages 3 (29–57 ka) and possibly 4 (57–71 ka)—at much greater densities than their counterparts in Australia for much of the past 50 ka—or that smaller hunter-gatherer groups banded together to make one or more migrations to Sahul. In either case, our results imply that modern human populations at that time

were sufficiently socially integrated<sup>34</sup> to be able to achieve the construction and successful voyaging of multiple ocean-going vessels<sup>27</sup>. This further suggests cognitive ability and planning, and probably deliberate migration given the numbers of people involved<sup>25,27</sup>.

Unlike the mostly genetics-based estimates of founding effective population size ( $n_e$ ) that cannot easily discern an associated census population size ( $n_c$ ), our estimates provide a value of  $n_c$  that does not necessarily imply random breeding among all individuals alive. This is because we indirectly accounted for potential inbreeding depression that could arise from non-random breeding by including a catastrophic mortality function that scales with generation time<sup>35</sup>. This added stochasticity thus more closely aligns with the ecological reality of a population constrained not only by environmental variability, but also by demographic and genetic stochasticity. Furthermore, genetics-based estimates of  $n_c$  cannot typically identify fine-scale details of multiple arrival events over the period of several human generations. Thus, our resampling approach also provides the unique minimum interval over which successive human arrivals could have occurred. This does not necessarily imply that arrivals of small groups of humans occurred over the 700- to 900-year (that is, 25- to 32-generation) window we estimated; rather, it merely indicates that the extinction probability remained low within this window. This does not therefore preclude the occurrence of larger and more frequent introduction pulses over longer time frames.

Of course, our model predictions do rely on several unmeasurable parameters, not least of which are the types of survival and fertility schedules experienced by the first humans to enter Sahul over 50–65 ka. Indeed, our global sensitivity analysis showed that our results are most sensitive to variation in the underlying patterns of initial and environmentally stochastic survival probabilities estimated with the Siler hazard model, as well as total fertility (Supplementary Fig. 3). Nonetheless, our adoption of ‘average’ hunter-gatherer demographic rates appears reasonable and probably does not overestimate infant survival<sup>31</sup>. We also assumed that carrying capacity was proportional to net primary productivity<sup>36</sup>, although it is also plausible that ancient humans struck a compromise between high productivity and ease of passage and/or visibility to hunt prey by tending towards ecotones of mid-range productivity<sup>37</sup>. Had the ecological conditions at the time of first arrival favoured higher vital rates, the true population size might have been larger than our estimates suggest; however, we are concerned here solely with estimating the minimum viable population size derived from conservative, yet realistic, demographic parameters.

More importantly, assessments of relative carrying capacity appeared to have only weak effects on our model predictions, particularly given the near-identical form of the quasi-extinction/founding-population-size curve for the introduction windows of 65–55 and 60–50 ka (Fig. 1), as well as the low influence of the density-feedback survival modifier ( $S_{\text{mod}}$ ) and nadir population density ( $D_{\text{min}}$ ) identified in the global sensitivity analysis (Supplementary Fig. 3). Thus, the specific choice of carrying capacity (expressed as the total number of humans permitted to occupy the landscape) and the arrival window per se have little bearing on our conclusions. This outcome holds even if carrying capacities were, in fact, higher than we assumed because of potentially higher prey availability at initial arrival relative to later periods when many megafauna species were no longer present<sup>1,2</sup> (although regional extirpation times of megafauna are still highly uncertain, especially for northern Australia).

Furthermore, our estimate of ~1,300 minimum founding individuals arriving within 25–32 generations agrees well with genetics-based estimates of total effective population size. For example, there are  $n_e$  estimates of populations as small as 170–230 Maori women based on mitochondrial DNA for the peopling of New Zealand<sup>38</sup>, as few as 70 individuals based on mitochondrial and Y-chromosome DNA for the peopling of the New World from Asia<sup>39</sup>, and

150 female Yakuts in north-eastern Siberia based on mitochondrial DNA<sup>40</sup>. While the true  $n_e:n_c$  ratio for each of these populations is unknown, if we assume an average of ~0.10 based on a multispecies assessment<sup>41</sup>, the previously cited values of  $n_c$  would equate to an  $n_e$  of 700–4,600 individuals; the resemblance to our demographically based estimates is therefore striking. Previous studies of Aboriginal Australian DNA posited that at least 36–50 founding females were required to establish known Australian mitochondrial diversity (estimated from 4–5 founding females for each of the 9–10 haplogroups)<sup>25</sup>. This range is likely to be a conservative minimum boundary, as founding populations with limited female diversity (<10 per haplogroup) have little chance of survival<sup>42</sup>. Again, assuming that  $n_e:n_c=0.1$  gives a value of  $n_c$  ranging from 720 to 1,000, our results are not at odds with this argument, although our model conservatively suggests that the minimum number of females per haplogroup would probably have been higher. However, such estimates assume that known Australian haplogroups today represent the total mitochondrial diversity present during the period of first arrival, and this assumption is likely to be inaccurate.

In summary, our demographic models quantifying the ecologically plausible demographic context of the first humans to people Australia now allow for exploration of other questions regarding human adaptations and technological developments during this period, which could have assisted in successfully populating Sahul. Possibly driven in part by the amalgamation of people in Sunda and Wallacea at this time, more research describing the antecedent conditions in those regions would assist greatly in describing the source population(s), and possibly reveal the impetus for subsequent directed migrations<sup>22–27</sup> to Sahul.

## Methods

**Demographic rates.** Our first requirement was to estimate realistic demographic rates (survival, fertility and longevity) for ancient Australians, to parameterize an age-structured model. For survival, we used the five-parameter Siler hazard model<sup>30</sup> to estimate the age ( $x$ )-specific proportion of surviving individuals ( $l_x$ ), which incorporates survival schedules for three stages: immature, mature and senescent individuals within the population:

$$l_x = e^{\left(\frac{-a_1}{b_1}\right)(1-e^{-b_1x})} e^{-a_2x} e^{\left(\frac{a_3}{b_3}\right)(1-e^{-b_3x})} \quad (1)$$

where  $a_1$  is the initial infant mortality (also described elsewhere as  $\alpha_1$ ),  $b_1$  is the rate of mortality decline in immatures,  $a_2$  is the age-independent mortality due to environmental influence,  $a_3$  is the initial adult mortality and  $b_3$  is the rate of mortality increase (senescence). We used the average ‘hunter-gatherer’ parameter estimates from Gurven and Kaplan<sup>31</sup> (based on modern populations and compared with palaeo-demography) to construct  $l_x$  ( $a_1 = 0.422$ ;  $b_1 = 1.131$ ;  $a_2 = 0.013$ ;  $a_3 = 0.000147$ ;  $b_3 = 0.086$ ), then calculated the age-specific survival ( $S_x$ ) for life tables (Supplementary Fig. 1) as:

$$S_x = 1 - \frac{(l_x - l_{x+1})}{l_x} \quad (2)$$

From the  $l_x$  vector, we also calculated the age-specific life expectancy ( $e_x$ ):

$$e_x = \sum_{i=1}^n \left( \frac{l_x + l_{x+1}}{2} \right)_i$$

For fertility, we first estimated a fertility schedule based on age at primiparity estimates for 22 modern hunter-gatherer groups<sup>43</sup>, taking the average and 95% confidence interval (CI) limits as indicators of the onset of reproduction in such societies. This gave a mean age of 19 years for primiparity among women (95% CI: 16–24). Further evidence of reproductive senescence and menopause in hunter-gatherer women suggests that hunter-gatherer societies included many women beyond their fertile years<sup>44–46</sup>. Thus, the onset of reproduction and the implied fertility decline compare well with the global average fertility schedule of modern *Homo sapiens*<sup>47</sup>. For total fertility ( $F$ ), we used the value of 4.69 births (that is, 2.35 daughters) for the !Kung hunter-gatherer society<sup>48</sup> (Supplementary Fig. 2).

**Age-structured (Leslie) matrix model.** From these estimated demographic rates, we constructed a prebreeding, 81 ( $i$ )  $\times$  81 ( $j$ ) element (representing ages from 0–80 years old), Leslie projection matrix ( $M$ ) for females only (males are demographically irrelevant in this context assuming equal sex ratios), multiplying

a population vector  $\mathbf{n}$  to estimate the total population size at each forecast time step<sup>49</sup>. Thus, we used a longevity ( $\omega$ ) of 80 years based on cross-cultural examination of hunter-gatherer societies<sup>51</sup>, which was itself founded on the modal adult death of about 70 years. Fertilities ( $m_x$ ) occupied the first row of the matrix, survival probabilities ( $S_x$ ) occupied the subdiagonal, and we set the final diagonal transition probability ( $M_{ii}$ ) to zero. We projected the  $M\mathbf{n}$  combinations for each iteration of the simulation (see below) to obtain the yearly total population size.

**Carrying capacity.** In the absence of measured compensatory density-feedback mechanisms for ancient humans, we used a hypothetical reduction in the survival vector by constructing a theoretical carrying capacity ( $K$ ) built from a hindcasted estimate of net primary production based on the LOVECLIM climate reconstruction<sup>50</sup>. LOVECLIM is a three-dimensional Earth system model of intermediate complexity<sup>51</sup> (that is, its spatial resolution is coarser than that of state-of-the-art general circulation models, and its representation of physical processes is simpler) that has been validated extensively for the past few glacial cycles and for many regions of the world<sup>52–55</sup>. LOVECLIM includes representations of the atmosphere, ocean and sea ice, land surface (including vegetation), ice sheets, icebergs and the carbon cycle, and produces climates over the past 120 ka in 1,000-year snapshots downscaled (using a bilinear interpolation)<sup>56,57</sup> at a spatial resolution of  $1^\circ \times 1^\circ$ . For each grid cell and each 1,000-year snapshot, we extracted the mean annual temperature, mean annual precipitation, freshwater availability (that is, evapotranspiration minus precipitation), bottom (soil) moisture, desert fraction and net primary production<sup>52</sup>. The candidate  $K$  output variables (freshwater availability, bottom moisture, desert fraction and net primary production) for northern Australia (see below) were highly correlated (Spearman's  $|\rho| \geq 0.842$ ; Supplementary Table 1), so we chose net primary production ( $\text{kgC m}^{-2} \text{yr}^{-1}$ ) as the comprehensive indicator of relative carrying capacity through time. Indeed, regional carrying capacity is correlated with net primary production for many species, including humans<sup>58–62</sup>. To focus on the region of interest, we took all Sahul (Australia, New Guinea and Tasmania)  $1^\circ \times 1^\circ$  grid cells from the equator ( $0^\circ$ ) to  $14^\circ$  S latitude to represent 'northern' Sahul (including New Guinea, most of the top end of the Northern Territory and Cape York Peninsula) (Fig. 2a), and calculated the 25th and 75th percentiles for net primary production across this region; from within this quartile range, we stochastically sampled annual net primary production per projection iteration (see below).

To translate net primary production into a carrying capacity expressed in units of humans the landscape was capable of supporting, we used data derived from archaeological sites and the assumption of a putative population low (nadir) that occurred during and immediately after the Last Glacial Maximum (23–18 ka)<sup>63–66</sup>, when conditions were cooler than today, and much (but not all<sup>67–70</sup>) of the continent was drier<sup>71–78</sup>. Demographic reconstructions based on the spatial distribution of dated archaeological sites suggest that up to 80% of Australia could have been abandoned or experienced reduced occupation at some point during this interval<sup>65</sup>, or at least a major spatial thinning of populations (perhaps as much as 60%) during this period<sup>9</sup>. From these demographic reconstructions, we set the baseline population size at the Last Glacial Maximum at 47,000 people continent wide<sup>92</sup>. This figure is based on an estimated area of habitable land<sup>79</sup> for Sahul of 9.4 million  $\text{km}^2$  and a population density<sup>65</sup> of 0.005 individuals  $\text{km}^{-2}$ , which is similar to historical estimates of population densities for Australian deserts<sup>80</sup>. We recognize that inferences of past population size are subject to many uncertainties<sup>66,81–84</sup> and note accordingly that our model results are not critically dependent on the above input values (see the global sensitivity analysis and Supplementary Fig. 3).

From this putative population low at or around the Last Glacial Maximum, we back-tracked to the window of first arrival to estimate a relative carrying capacity for this period. We then scaled the relative net primary production curve by first adding the absolute minimum 25th percentile to each annual value, and then dividing by the maximum median value. To these scaled annual net primary production values, we multiplied by 47,000 people to provide an annual  $K$  in units of individual people (Fig. 2b). For the founding period of interest (65–50 ka), this translates into a minimum  $K$  of 69,230–111,329 individuals (25th–75th percentile limits) at 55 ka, and a maximum  $K$  of 82,297–158,645 individuals (25th–75th percentile limits) from 63–62 ka (Fig. 2c). We also reproduced the analysis with a starting window between 60 and 50 ka, assuming instead a later date of first arrival (see Results). It is important to understand that the precise timing of the putative population nadir is irrelevant from the perspective of the mathematical reconstruction of the  $K$  series, as long as a nadir occurred at some point after first arrival. Also, the specific  $K$  conditions at time of first arrival had little effect on our model outputs (see Results).

**Compensatory density feedback.** When the projected population exceeded the resampled net primary production  $K$  in person units that year, we multiplied the  $\beta$ -resampled survival vector (see below) by a multiplier of 0.98 ( $S_{\text{mod}}$ ) to impose a compensatory feedback mechanism. This is because the base  $M$  matrix had a low dominant eigenvalue (that is, rate of population change; see Results), so this compensatory density-feedback mechanism amounts to a 2% drop in average survival each time total abundance exceeded that time step's sampled  $K$  value. This acted to keep the projections from growing exponentially over the 100 human generations.

**Catastrophic mortality events.** We also included a catastrophic die-off function in the simulations to account for the probability of catastrophic mortality events ( $C$ ) scaling to generation length among vertebrates<sup>55</sup>:

$$C = \frac{P_C}{G}$$

where  $P_C$  is the probability of catastrophe (set at 0.14)<sup>55</sup> and  $G$  is the mean generation time (27.6 years), as calculated from the deterministic matrix  $M$ <sup>49</sup>. Once invoked at probability  $C$  for any iteration of the model (see below), we halved the survival vector to induce a 50% mortality ( $d$ ) event for that year<sup>85</sup>. This was based on the definition of a catastrophe as "... any 1 year peak-to-trough decline in estimated numbers of 50% or greater"<sup>35</sup>.

**Stochastic projections.** We conservatively sampled the start date for each of 10,000 projection iterations using a stochastic uniform sampler between 65 and 50 ka (we aimed to use the full uncertainty of  $K$  during the approximate window of first arrival). We thus had a different, randomly selected start year for the 100 generations projected into the future (that is, from 65–50 ka towards the present), based on the stochastically sampled  $M$  matrix elements. Here, we defined a function to estimate the shape parameters of a  $\beta$  function, and then randomly  $\beta$ -resampled each element of the survival vector for each year of the projection (assuming an arbitrary  $\sigma_3 = 5\%$  standard deviation on survival probability). For the fertility vector, we used a random Gaussian resampler for the total (female) fertility  $F$  described above, based also on an arbitrary 5% standard deviation.

**Founding population size.** We applied a starting population size from 50–1,000 females in increments of 50, and calculated the probability of quasi-extinction as the number of iterations per founding population size, where at least one projected annual total population size fell below a quasi-extinction threshold ( $Q$ ) of 50 individuals (that is, 25 females, assuming equal sex ratios). This was based on the minimum size below which a population cannot avoid inbreeding depression (although it could be twice as high as this<sup>86</sup>, so our approach was conservative).

To estimate a realistic extinction risk, we must borrow from the ecological concept of minimum viable population size<sup>87</sup>. Here, there is a rising consensus that several thousand individuals are normally required to avoid inbreeding depression, thereby maintaining evolutionary potential<sup>86</sup>, and thus avoiding extinction<sup>88</sup>. This is because non-random breeding generally equates to lower effective population sizes ( $n_e$ ) than census population sizes ( $n_c$ )<sup>41</sup>. In the case of founding *Homo* species populations, various population genetic approaches (in some instances combined with archaeological evidence<sup>9</sup>) have estimated minimum founder population sizes from 80 to several thousand effective individuals<sup>9,39,89–93</sup>. However, the relationship between  $n_e$  and  $n_c$  is complex and variable<sup>41</sup>, depending in part on the time frame over which the data are collected and measured<sup>94</sup>. Even with the current lack of reliable estimates of  $n_e$  for the first people to arrive in Australia, the uncertainty associated with  $n_e:n_c$  ratios means that another approach is required to estimate both the probable initial population size of founding humans arriving over 50 ka and the period during which these people probably arrived in Australia and became a genetically interacting and viable founding population.

However, this approach assumes an instantaneous arrival of the entire founding population in year 1, which is probably an unrealistic representation of the more likely sequence of multiple arrivals of smaller groups over the entire founding 'interval'. To estimate the frequency of smaller introduction events that maintained a low probability of extinction, we resampled 10,000 times the range of the minimum viable population defined in the previous step (that is, the minimum number of total founders maintaining a probability of quasi-extinction of ~0.1). We first assumed that each introduction event represented one-tenth of the total founding population, but that they were spread out by an incrementing interval of decades. Thus, the first introduction frequency was every 10 years (that is, one-tenth of the minimum viable founding population arriving every 10 years over one century), the second was every 20 years (one-tenth every 20 years over two centuries), and so on until a frequency of 300 years (that is, one-tenth arriving every 300 years over 3,000 years) (Fig. 1c,d). The resulting frequency-quasi-extinction probability relationship thus indicates the frequency at which one-tenth of the minimum founding population is required to raise the probability of extinction beyond the level of ~0.1 established in the first step.

However, the reality of an even frequency of identical arriving population sizes is also unlikely, so we added complexity to our model (thus increasing realism) by randomly resampling 10,000 times both the number of introduction events and the frequency between events, such that the latter averaged an incrementing range of decades between events (as above). Here, we randomly resampled the initial introduction event as a random uniform number between 25 females (quasi-extinction threshold) and one-half of the minimum founding population established in the first step. We then randomly resampled the following introduction-population sizes from the remaining number of individuals up to the minimum total founding population size, until we reached the cumulative minimum founding population size. We used a random Gaussian sampler of the same sequence as in the previous step, assuming a 10% standard deviation. Thus, the first frequency was an introduction interval resampled with a mean of 10 years

and a standard deviation of 1 year, the second was resampled with a mean of 20 years and a standard deviation of 2 years, and so on up to a mean of 200 years between introductions (and the associated 20-year standard deviation).

**Global sensitivity analysis.** We designed a 'global' sensitivity analysis to provide robust sensitivity measures of the probability of quasi-extinction to variation in the underlying parameters of our stochastic model<sup>55,56</sup>. We applied a Latin hypercube-sampling protocol<sup>56</sup> of the parameter space assuming a founding population size of 700 females projected over 100 generations. We sampled 12 parameters from a uniform distribution as follows: (1–5) all five parameters used to calculate the Siler hazard model for age-specific survival:  $a_1$  (varying from 0.3–0.5),  $b_1$  (1.0–2.0),  $a_2$  (0.010–0.015),  $a_3$  ( $1.323 \times 10^{-4}$ – $1.617 \times 10^{-4}$ ) and  $b_3$  (0.060–0.095); (6) standard deviation of survival ( $\sigma_s$ ) for stochastic resampling (0.025–0.100); (7) density-feedback survival modifier  $S_{\text{mod}}$  (0.95–0.99); (8) total fertility  $F$  (2.1105–2.5795); (9) quasi-extinction threshold  $Q$  (13–75 females); (10) probability of catastrophe  $P_C$  (0.1–0.2); (11) intensity of catastrophic die-offs  $d$  (0.25–0.75); and (12) nadir population density during/near the Last Glacial Maximum (0.0025–0.010 individuals km<sup>-2</sup>; that is, from half to double the 0.005 value assumed in the model based on archaeological data<sup>65,79</sup>). To sample using the Latin hypercube protocol, we ran the simulation for 100 iterations, with 1,000 samples from the parameter space. To test the effect of the parameter values on  $P$  (quasi-extinction), we used a boosted-regression tree<sup>97</sup> emulator with the function `gbm.step`<sup>98</sup> in the `dismo` R library, setting the error distribution family as Gaussian, the bag fraction to 0.75, the learning rate to 0.01, the tolerance to 0.0001 and the tree complexity to 2 (first-order interactions only). To assess the relative contribution of each sampled parameter to  $P$  (quasi-extinction), we present the boosted-regression tree metrics of relative influence<sup>96</sup>.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

All data are available for download at [github.com/cjabradshaw/SahulHuman](https://github.com/cjabradshaw/SahulHuman).

### Code availability

All R code is available for download at [github.com/cjabradshaw/SahulHuman](https://github.com/cjabradshaw/SahulHuman).

Received: 3 December 2018; Accepted: 15 April 2019;

Published online: 17 June 2019

### References

- Saltré, F. et al. Climate change not to blame for Late Quaternary megafauna extinctions in Australia. *Nat. Comm.* **7**, 10511 (2016).
- Johnson, C. N. et al. What caused extinction of the Pleistocene megafauna of Sahul? *Proc. R. Soc. Lond. B* **283**, 20152399 (2016).
- O'Connell, J. F. et al. When did *Homo sapiens* first reach Southeast Asia and Sahul? *Proc. Natl Acad. Sci. USA* **115**, 8482–8490 (2018).
- Birdsell, J. B. Some population problems involving Pleistocene man. *Cold Spring Harb. Symp. Quant. Biol.* **22**, 47–69 (1957).
- McArthur, N. Computer simulations of small populations. *Aust. Archaeol.* **4**, 53–57 (1976).
- Allen, J. & O'Connell, J. F. In *Islands of Inquiry: Colonisation, Seafaring and the Archaeology of Maritime Landscapes, Terra Australis* Vol. 29 (eds Clark, G., Leach, F. & O'Connell, S.) 31–46 (ANU E Press, 2008).
- O'Connell, J. F. & Allen, J. The restaurant at the end of the universe: modelling the colonisation of Sahul. *Aust. Archaeol.* **74**, 5–17 (2012).
- Rasmussen, M. et al. An Aboriginal Australian genome reveals separate human dispersals into Asia. *Science* **334**, 94–98 (2011).
- Williams, A. N. A new population curve for prehistoric Australia. *Proc. R. Soc. Lond. B* **280**, 20130486 (2013).
- Malaspina, A.-S. et al. A genomic history of Aboriginal Australia. *Nature* **538**, 207–214 (2016).
- Clarkson, C. et al. Human occupation of northern Australia by 65,000 years ago. *Nature* **547**, 306–310 (2017).
- Clarkson, C. et al. Reply to comments on Clarkson et al. (2017) 'Human occupation of northern Australia by 65,000 years ago'. *Aust. Archaeol.* **84**, 84–89 (2018).
- Roberts, R. G. et al. The human colonisation of Australia: optical dates of 53,000 and 60,000 years bracket human arrival at Deaf Adder Gorge, Northern Territory. *Quat. Sci. Rev.* **13**, 575–583 (1994).
- Turney, C. S. M. et al. Early human occupation at Devil's Lair, southwestern Australia 50,000 years ago. *Quat. Res.* **55**, 3–13 (2011).
- Bowler, J. M. et al. New ages for human occupation and climatic change at Lake Mungo, Australia. *Nature* **421**, 837–840 (2003).
- Wood, R. et al. Towards an accurate and precise chronology for the colonization of Australia: the example of Riwi, Kimberley, Western Australia. *PLoS ONE* **11**, e0160123 (2016).
- Hamm, G. et al. Cultural innovation and megafauna interaction in the early settlement of arid Australia. *Nature* **539**, 280–283 (2016).
- Veth, P. et al. Early human occupation of a maritime desert, Barrow Island, north-west Australia. *Quat. Sci. Rev.* **168**, 19–29 (2017).
- Delannoy, J.-J. et al. in *The Archaeology of Rock Art in Western Arnhem Land, Australia, Terra Australis* Vol. 47 (eds David, B., Taçon, P. S. C., Delannoy, J.-J. & Geneste, J.-M.) 197–243 (ANU Press, 2017).
- Maloney, T., O'Connor, S., Wood, R., Aplin, K. & Balme, J. Carpenters Gap 1: a 47,000 year old record of indigenous adaptation and innovation. *Quat. Sci. Rev.* **191**, 204–228 (2018).
- McDonald, J. et al. *Karnatukul* (Serpent's Glen): a new chronology for the oldest site in Australia's Western Desert. *PLoS ONE* **13**, e0202511 (2018).
- Kealy, S., Louys, J. & O'Connor, S. Islands under the sea: a review of early modern human dispersal routes and migration hypotheses through Wallacea. *J. Isl. Coast. Archaeol.* **11**, 364–384 (2016).
- Kealy, S., Louys, J. & O'Connor, S. Reconstructing palaeogeography and inter-island visibility in the Wallacean Archipelago during the likely period of Sahul colonization, 65–45000 years ago. *Archaeol. Prospect.* **24**, 259–272 (2017).
- Norman, K. et al. An early colonisation pathway into northwestern Australia 70–60,000 years ago. *Quat. Sci. Rev.* **180**, 229–239 (2018).
- Bird, M. I. et al. Palaeogeography and voyage modeling indicates early human colonization of Australia was likely from Timor-Roti. *Quat. Sci. Rev.* **191**, 431–439 (2018).
- Kealy, S., Louys, J. & O'Connor, S. Least-cost pathway models indicate northern human dispersal from Sunda to Sahul. *J. Hum. Evol.* **125**, 59–70 (2018).
- Bird, M. I. et al. Early human settlement of Sahul was not an accident. *Sci. Rep.* <https://doi.org/10.1038/s41598-019-42946-9> (2019).
- Nagle, N. et al. Aboriginal Australian mitochondrial genome variation—an increased understanding of population antiquity and diversity. *Sci. Rep.* **7**, 43041 (2017).
- Tobler, R. et al. Aboriginal mitogenomes reveal 50,000 years of regionalism in Australia. *Nature* **544**, 180–184 (2017).
- Siler, W. A competing-risk model for animal mortality. *Ecology* **60**, 750–757 (1979).
- Gurven, M. & Kaplan, H. Longevity among hunter-gatherers: a cross-cultural examination. *Pop. Dev. Rev.* **33**, 321–365 (2007).
- Fenner, J. N. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am. J. Phys. Anthropol.* **128**, 415–423 (2005).
- Gould, R. A. *Puntutjarpa Rockshelter and the Australian Desert Culture* Anthropological Papers of the American Museum of Natural History Vol. 54 (American Museum of Natural History, 1977).
- Durkheim, É. *The Division of Labour in Society* (Macmillan, 1984).
- Reed, D. H., O'Grady, J. J., Ballou, J. D. & Frankham, R. The frequency and severity of catastrophic die-offs in vertebrates. *Anim. Conserv.* **6**, 109–114 (2003).
- Tallavaara, M., Eronen, J. T. & Luoto, M. Productivity, biodiversity, and pathogens influence the global hunter-gatherer population density. *Proc. Natl Acad. Sci. USA* **115**, 1232–1237 (2018).
- Finlayson, C. et al. The *Homo* habitat niche: using the avian fossil record to depict ecological characteristics of Palaeolithic Eurasian hominins. *Quat. Sci. Rev.* **30**, 1525–1532 (2011).
- Whyte, A. L. H., Marshall, S. J. & Chambers, G. K. Human evolution in Polynesia. *Hum. Biol.* **77**, 157–177 (2005).
- Hey, J. On the number of New World founders: a population genetic portrait of the peopling of the Americas. *PLoS Biol.* **3**, e193 (2005).
- Zlojutro, M. et al. Coalescent simulations of Yakut mtDNA variation suggest small founding population. *Am. J. Phys. Anthropol.* **139**, 474–482 (2009).
- Frankham, R. Effective population size/adult population size ratios in wildlife: a review. *Genet. Res.* **66**, 95–107 (1995).
- Moore, J. H. Evaluating five models of human colonization. *Am. Anthropol.* **103**, 395–408 (2001).
- Walker, R. et al. Growth rates and life histories in twenty-two small-scale societies. *Am. J. Hum. Biol.* **18**, 295–311 (2006).
- Hawkes, K., Smith, K. R. & Robson, S. L. Mortality and fertility rates in humans and chimpanzees: how within-species variation complicates cross-species comparisons. *Am. J. Hum. Biol.* **21**, 578–586 (2009).
- Hawkes, K. & Coxworth, J. E. Grandmothers and the evolution of human longevity: a review of findings and future directions. *Evol. Anthropol.* **22**, 294–302 (2013).
- Blurton Jones, N. G., Hawkes, K. & O'Connell, J. F. Antiquity of postreproductive life: are there modern impacts on hunter-gatherer postreproductive life spans? *Am. J. Hum. Biol.* **14**, 184–205 (2002).
- Bradshaw, C. J. A. & Brook, B. W. Human population reduction is not a quick fix for environmental problems. *Proc. Natl Acad. Sci. USA* **111**, 16610–16615 (2014).
- Bentley, G. R. Hunter-gatherer energetics and fertility: a reassessment of the !Kung San. *Hum. Ecol.* **13**, 79–109 (1985).
- Caswell, H. *Matrix Population Models: Construction, Analysis, and Interpretation* 2nd edn (Sinauer Associates, 2001).

50. Goosse, H. et al. Description of the Earth system model of intermediate complexity LOVECLIM version 1.2. *Geosci. Mod. Dev.* **3**, 603–633 (2010).
51. Claussen, M. et al. Earth system models of intermediate complexity: closing the gap in the spectrum of climate system models. *Clim. Dynam.* **18**, 579–586 (2002).
52. Timmermann, A. & Friedrich, T. Late Pleistocene climate drivers of early human migration. *Nature* **538**, 92–95 (2016).
53. Friedrich, T., Timmermann, A., Tigchelaar, M., Elison Timm, O. & Ganopolski, A. Nonlinear climate sensitivity and its implications for future greenhouse warming. *Sci. Adv.* **2**, e1501923 (2016).
54. Tigchelaar, M., Timmermann, A., Pollard, D., Friedrich, T. & Heinemann, M. Local insolation changes enhance Antarctic interglacials: insights from an 800,000-year ice sheet simulation with transient climate forcing. *Earth Planet. Sci. Lett.* **495**, 69–78 (2018).
55. Stockhecke, M. et al. Millennial to orbital-scale variations of drought intensity in the Eastern Mediterranean. *Quat. Sci. Rev.* **133**, 77–95 (2016).
56. Lorenz, D. J., Nieto-Lugilde, D., Blois, J. L., Fitzpatrick, M. C. & Williams, J. W. Downscaled and debiased climate simulations for North America from 21,000 years ago to 2100 AD. *Sci. Data* **3**, 160048 (2016).
57. Wilby, R. L. & Wigley, T. M. L. Downscaling general circulation model output: a review of methods and limitations. *Prog. Phys. Geogr. Earth Environ.* **21**, 530–548 (1997).
58. Coe, M. J., Cumming, D. H. & Phillipson, J. Biomass and production of large African herbivores in relation to rainfall and primary production. *Oecologia* **22**, 341–354 (1976).
59. Krausmann, F. et al. Long-term trajectories of the human appropriation of net primary production: lessons from six national case studies. *Ecol. Econ.* **77**, 129–138 (2012).
60. Whittaker, R. H. & Likens, G. E. Primary production: the biosphere and man. *Hum. Ecol.* **1**, 357–369 (1973).
61. Phillipson, J. Rainfall, primary production and ‘carrying capacity’ of Tsavo National Park (East), Kenya. *Afr. J. Ecol.* **13**, 171–201 (1975).
62. Cao, M., Ma, S. & Han, C. Potential productivity and human carrying capacity of an agro-ecosystem: an analysis of food production potential of China. *Agric. Syst.* **47**, 387–414 (1995).
63. Williams, M. et al. Glacial and deglacial climatic patterns in Australia and surrounding regions from 35000 to 10000 years ago reconstructed from terrestrial and near-shore proxy data. *Quat. Sci. Rev.* **28**, 2398–2419 (2009).
64. Petherick, L. M., Moss, P. T. & McGowan, H. A. Climatic and environmental variability during the termination of the Last Glacial Stage in coastal eastern Australia: a review. *Aust. J. Earth Sci.* **58**, 563–577 (2011).
65. Williams, A. N., Ulm, S., Cook, A. R., Langley, M. C. & Collard, M. Human refugia in Australia during the Last Glacial Maximum and terminal Pleistocene: a geospatial analysis of the 25–12 ka Australian archaeological record. *J. Archaeol. Sci.* **40**, 4612–4625 (2013).
66. Williams, A. N. et al. A continental narrative: human settlement patterns and Australian climate change over the last 35,000 years. *Quat. Sci. Rev.* **123**, 91–112 (2015).
67. Ellerton, D., Shulmeister, J., Woodward, C. & Moss, P. Last Glacial Maximum and Last Glacial–Interglacial Transition pollen record from northern NSW, Australia: evidence for a humid late Last Glacial Maximum and dry deglaciation in parts of eastern Australia. *J. Quat. Sci.* **32**, 717–728 (2017).
68. Hesse, P. P. et al. Dramatic reduction in size of the lowland Macquarie River in response to Late Quaternary climate-driven hydrologic change. *Quat. Res.* **90**, 360–379 (2018).
69. Shulmeister, J., Kemp, J., Fitzsimmons, K. E. & Gontz, A. Constant wind regimes during the Last Glacial Maximum and Early Holocene: evidence from Little Llangothlin Lagoon, New England Tablelands, eastern Australia. *Clim. Past* **12**, 1435–1444 (2016).
70. Mueller, D. et al. Revisiting an arid LGM using fluvial archives: a luminescence chronology for palaeochannels of the Murrumbidgee River, south-eastern Australia. *J. Quat. Sci.* **33**, 777–793 (2018).
71. Hope, G. et al. History of vegetation and habitat change in the Austral-Asian region. *Quat. Int.* **118–119**, 103–126 (2004).
72. Johnson, B. J. et al. 65,000 years of vegetation change in central Australia and the Australian summer monsoon. *Science* **284**, 1150–1152 (1999).
73. Fitzsimmons, K. E. et al. Late Quaternary palaeoenvironmental change in the Australian drylands. *Quat. Sci. Rev.* **74**, 78–96 (2013).
74. Barrows, T. T., Stone, J. O. & Fifield, L. K. Exposure ages for Pleistocene periglacial deposits in Australia. *Quat. Sci. Rev.* **23**, 697–708 (2004).
75. Barrows, T. T., Stone, J. O., Fifield, L. K. & Cresswell, R. G. The timing of the Last Glacial Maximum in Australia. *Quat. Sci. Rev.* **21**, 159–173 (2002).
76. Barrows, T. T., Stone, J. O., Fifield, L. K. & Cresswell, R. G. Late Pleistocene glaciation of the Kosciuszko Massif, Snowy Mountains, Australia. *Quat. Res.* **55**, 179–189 (2001).
77. Reeves, J. M. et al. Climate variability over the last 35,000 years recorded in marine and terrestrial archives in the Australian region: an OZ-INTIMATE compilation. *Quat. Sci. Rev.* **74**, 21–34 (2013).
78. Petherick, L. et al. Climatic records over the past 30 ka from temperate Australia—a synthesis from the Oz-INTIMATE workgroup. *Quat. Sci. Rev.* **74**, 58–77 (2013).
79. Gautney, J. R. & Holliday, T. W. New estimations of habitable land area and human population size at the Last Glacial Maximum. *J. Archaeol. Sci.* **58**, 103–112 (2015).
80. Smith, M. *The Archaeology of Australia's Deserts* (Cambridge Univ. Press, 2013).
81. Attenbrow, V. & Hiscock, P. Dates and demography: are radiometric dates a robust proxy for long-term prehistoric demographic change? *Archaeol. Oceania* **50**, 30–36 (2015).
82. Hiscock, P. & Attenbrow, V. Dates and demography? The need for caution in using radiometric dates as a robust proxy for prehistoric population change. *Archaeol. Oceania* **51**, 218–219 (2016).
83. Smith, M. The use of summed-probability plots of radiocarbon data in archaeology. *Archaeol. Oceania* **51**, 214–215 (2016).
84. Williams, A. N. & Ulm, S. Radiometric dates are a robust proxy for long-term demographic change: a comment on Attenbrow and Hiscock (2015). *Archaeol. Oceania* **51**, 216–217 (2016).
85. Bradshaw, C. J. A. et al. More analytical bite in estimating targets for shark harvest. *Mar. Ecol. Prog. Ser.* **488**, 221–232 (2013).
86. Frankham, R., Bradshaw, C. J. A. & Brook, B. W. Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biol. Conserv.* **170**, 56–63 (2014).
87. Traill, L. W., Bradshaw, C. J. A. & Brook, B. W. Minimum viable population size: a meta-analysis of 30 years of published estimates. *Biol. Conserv.* **139**, 159–166 (2007).
88. Traill, L. W., Brook, B. W., Frankham, R. & Bradshaw, C. J. A. Pragmatic population viability targets in a rapidly changing world. *Biol. Conserv.* **143**, 28–34 (2010).
89. Harpending, H. C. et al. Genetic traces of ancient demography. *Proc. Natl Acad. Sci. USA* **95**, 1961–1967 (1998).
90. Murray-McIntosh, R. P., Scrimshaw, B. J., Hatfield, P. J. & Penny, D. Testing migration patterns and estimating founding population size in Polynesia by using human mtDNA sequences. *Proc. Natl Acad. Sci. USA* **95**, 9047–9052 (1998).
91. Tenesa, A. et al. Recent human effective population size estimated from linkage disequilibrium. *Genome Res.* **17**, 520–526 (2007).
92. Liu, H., Prugnolle, F., Manica, A. & Balloux, F. A geographically explicit genetic model of worldwide human-settlement history. *Am. J. Hum. Gen.* **79**, 230–237 (2006).
93. Zollner, P. A. & Lima, S. L. Search strategies for landscape-level interpatch movements. *Ecology* **80**, 1019–1030 (1999).
94. Eller, E., Hawks, J. & Relethford, J. H. Local extinction and recolonization, species effective population size, and modern human origins. *Hum. Biol.* **81**, 805–824 (2009).
95. Wainwright, H. M., Finsterle, S., Jung, Y., Zhou, Q. & Birkholzer, J. T. Making sense of global sensitivity analyses. *Comput. Geosci.* **65**, 84–94 (2014).
96. Prowse, T. A. A. et al. An efficient protocol for the global sensitivity analysis of stochastic ecological models. *Ecosphere* **7**, e01238 (2016).
97. Elith, J., Leathwick, J. R. & Hastie, T. A working guide to boosted regression trees. *J. Anim. Ecol.* **77**, 802–813 (2008).
98. Hijmans, R. J., Phillips, S., Leathwick, J. & Elith, J. dismo: Species distribution modeling. R package version 1.1-4 <https://cran.r-project.org/web/packages/dismo/index.html> (2017).

## Acknowledgements

This study was supported by the Australian Research Council through a Centre of Excellence grant (CE170100015) to R.G.R., S.U., M.I.B., Z.J., C.J.A.B. and L.S.W., fellowships to S.U. (FT120100656), M.I.B. (FL140100044), R.G.R. (FL130100116), Z.J. (FT150100138) and L.S.W. (FT180100407), and an Australian Government Research Training Program Award to K.N.

## Author contributions

C.J.A.B. and F.S. designed the research. C.J.A.B. performed the analysis and sourced the data. C.J.A.B., F.S., S.U., A.N.W. and M.I.B. wrote the paper. All other co-authors contributed substantially to developing the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41559-019-0902-6>.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Correspondence and requests for materials** should be addressed to C.J.A.B.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- |                                     |                                     |  |
|-------------------------------------|-------------------------------------|--|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A description of all covariates tested   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated   |

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

- |                 |   |
|-----------------|---|
| Data collection | All code was written in the R Project for Statistical Computing ( <a href="http://www.r-project.org">www.r-project.org</a> ), and is available at <a href="https://github.com/cjabradshaw/SahulHuman">https://github.com/cjabradshaw/SahulHuman</a> |
| Data analysis   | All code was written in the R Project for Statistical Computing ( <a href="http://www.r-project.org">www.r-project.org</a> ), and is available at <a href="https://github.com/cjabradshaw/SahulHuman">https://github.com/cjabradshaw/SahulHuman</a> |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data and R code are available for download at [github.com/cjabradshaw/SahulHuman](https://github.com/cjabradshaw/SahulHuman)

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences



## Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We designed a demographic model for ancient humans entering Sahul based on estimates of fertility, longevity, and survival, and landscape carrying capacities based on hindcasted net primary production estimates from general circulation models (LOVECLIM).
Research sample	Modelling study only; no actual specimens were sampled specifically in this study (i.e., published data only).
Sampling strategy	'sampling' was done stochastically to achieve Gaussian-like behaviour (10,000 iterations per incrementing scenario value)
Data collection	No data were collected as such; all data derived from published sources or derived via modelling
Timing and spatial scale	approximately 60,000 to 45,000 years ago; spatial scale = northern Australia (10 to 18 degrees South latitude)
Data exclusions	No data excluded
Reproducibility	We estimated probabilities based on stochastic resampling of the full parameter uncertainties, for which an accompanying global sensitivity analysis is presented.
Randomization	Full random resampling according to various distributions (Gaussian, uniform, beta).
Blinding	NA
Did the study involve field work?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Included in the study
<input type="checkbox"/>	<input type="checkbox"/> Antibodies
<input type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input type="checkbox"/> Human research participants
<input type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Included in the study
<input type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

### Antibodies

Antibodies used	Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.
Validation	Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.

### Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	State the source of each cell line used.
Authentication	Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.
Mycoplasma contamination	Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	Name any commonly misidentified cell lines used in the study and provide a rationale for their use.

## Palaeontology

Specimen provenance *Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).*

Specimen deposition *Indicate where the specimens have been deposited to permit free access by other researchers.*

Dating methods *If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals *For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.*

Wild animals *Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*

Field-collected samples *For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*

Ethics oversight *Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics *Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*

Recruitment *Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*

Ethics oversight *Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration *Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*

Study protocol *Note where the full trial protocol can be accessed OR if not available, explain why.*

Data collection *Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*

Outcomes *Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

## ChIP-seq

### Data deposition

Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).

Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, May remain private before publication. provide a link to the deposited data.*

Files in database submission *Provide a list of all files available in the database submission.*

Genome browser session (e.g. [UCSC](#)) *Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

## Methodology

Replicates *Describe the experimental replicates, specifying number, type and replicate agreement.*

Sequencing depth *Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*

Antibodies *Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*

Peak calling parameters *Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*

Data quality *Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*

Software *Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

## Flow Cytometry

### Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

Sample preparation *Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*

Instrument *Identify the instrument used for data collection, specifying make and model number.*

Software *Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*

Cell population abundance *Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*

Gating strategy *Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

## Magnetic resonance imaging

### Experimental design

Design type *Indicate task or resting state; event-related or block design.*

Design specifications *Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*

Behavioral performance measures *State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

## Acquisition

Imaging type(s)

Field strength

Sequence & imaging parameters

Area of acquisition

Diffusion MRI  Used  Not used

## Preprocessing

Preprocessing software

Normalization

Normalization template

Noise and artifact removal

Volume censoring

## Statistical modeling & inference

Model type and settings

Effect(s) tested

Specify type of analysis:  Whole brain  ROI-based  Both

Statistic type for inference (See [Eklund et al. 2016](#))

Correction

## Models & analysis

n/a | Involved in the study

Functional and/or effective connectivity

Graph analysis

Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Graph analysis

Multivariate modeling and predictive analysis