



OPEN

## Climate and demography drive 7000 years of dietary change in the Central Andes

Kurt M. Wilson<sup>1,2,3</sup>✉, Weston C. McCool<sup>1,3</sup>, Simon C. Brewer<sup>2,4</sup>, Nicole Zamora-Wilson<sup>1,3</sup>, Percy J. Schryver<sup>1,3</sup>, Roxanne Lois F. Lamson<sup>1,2,3</sup>, Ashlyn M. Huggard<sup>1,3</sup>, Joan Brenner Coltrain<sup>1,2,3</sup>, Daniel A. Contreras<sup>5</sup> & Brian F. Codding<sup>1,2,3</sup>

Explaining the factors that influence past dietary variation is critically important for understanding changes in subsistence, health, and status in past societies; yet systematic studies comparing possible driving factors remain scarce. Here we compile the largest dataset of past diet derived from stable isotope  $\delta^{13}\text{C}\text{‰}$  and  $\delta^{15}\text{N}\text{‰}$  values in the Americas to quantitatively evaluate the impact of 7000 years of climatic and demographic change on dietary variation in the Central Andes. Specifically, we couple paleoclimatic data from a general circulation model with estimates of relative past population inferred from archaeologically derived radiocarbon dates to assess the influence of climate and population on spatiotemporal dietary variation using an ensemble machine learning model capable of accounting for interactions among predictors. Results reveal that climate and population strongly predict diet (80% of  $\delta^{15}\text{N}\text{‰}$  and 66% of  $\delta^{13}\text{C}\text{‰}$ ) and that Central Andean diets correlate much more strongly with local climatic conditions than regional population size, indicating that the past 7000 years of dietary change was influenced more by climatic than socio-demographic processes. Visually, the temporal pattern suggests decreasing dietary variation across elevation zones during the Late Horizon, raising the possibility that sociopolitical factors overrode the influence of local climatic conditions on diet during that time. The overall findings and approach establish a general framework for understanding the influence of local climate and demography on dietary change across human history.

Identifying the long-term drivers of past spatiotemporal variation in individuals' diets remains one of the grand challenges in archaeology<sup>1</sup>, yet systematic studies that quantify the relative influence of competing factors remain limited. Two factors, climate and population, may be particularly influential in structuring dietary variation. Local climatic conditions, proxies for environments, structure resource availability and land use strategies, resulting in patterned adaptations across similar environments<sup>2–5</sup>. Changes in those local climatic conditions may constrain viable subsistence adaptations by limiting agricultural viability<sup>6–8</sup> or altering resource distributions, leading to dietary change. Demographically-driven resource competition may also drive major subsistence transitions<sup>9,10</sup>, including domestication<sup>11,12</sup> and intensification<sup>13</sup>, whose proximate causes are innovation or crises<sup>11,14</sup>. Additionally, increasing population can cause changes such as altering mobility patterns<sup>15</sup> and/or influence rising sociopolitical complexity<sup>16</sup>. Critically, the influence of both climatic<sup>17</sup> and demographic<sup>18</sup> factors goes beyond diet itself to impact health and longevity, as well as status and political complexity<sup>19</sup>. However, the interaction between climate and demography further complicates these dynamics: climatic change may structure population growth rates and carrying capacity<sup>20–22</sup> leading to changes in diet and/or dietary variation<sup>23</sup> or population change may drive people to alter their environments<sup>24–26</sup>, also potentially causing changes in diets and/or dietary variation. As a result, untangling the interactive drivers of climate and demography on spatiotemporal dietary variation and, therefore, its consequences for factors like health and status, remains difficult.

As a case study to systematically untangle and evaluate the influences of climatic and demographic change on diet we investigate the predictive power of both climate and population on individual diets over the past 7000 years in the Central Andes. Given the scarcity of long-term quantitative studies on diet in the region, first, we generate a database of 1965 archaeological individuals with stable isotope  $\delta^{15}\text{N}\text{‰}$  or  $\delta^{13}\text{C}\text{‰}$  bone collagen

<sup>1</sup>Department of Anthropology, University of Utah, 260 S. Central Campus Drive, Room 4625, Salt Lake City, UT 84112, USA. <sup>2</sup>Global Change and Sustainability Center, University of Utah, Salt Lake City, UT 84112, USA. <sup>3</sup>University of Utah Archaeological Center, University of Utah, Salt Lake City, UT 84112, USA. <sup>4</sup>Department of Geography, University of Utah, Salt Lake City, UT 84112, USA. <sup>5</sup>Department of Anthropology, University of Florida, Gainesville, FL 32603, USA. ✉email: kurt.wilson@utah.edu

values from which we reconstruct both the patterns of dietary change over space and time and the variation between individuals' diets. Due to the dramatic influence of elevation on environment<sup>27</sup> and subsistence strategy<sup>28</sup> in the Andes, we divide the study area into three elevational zones and classify each individual by location (see methods). The data reveal that Central Andean diets exhibit broad changes and vary significantly over time and space. To evaluate whether the observed variation results more from climatic factors or from population change, we analyze this spatially-explicit time series of individual diets using simulated paleoclimate derived from a general circulation model (GCM) and a population proxy generated from a dataset of 3957 archaeological radiocarbon dates. To address the analytic challenge posed by the likelihood that climatic and demographic change are linked<sup>22,29</sup>, we conduct the first quantitative inter-regional analysis of trans-Holocene diets in the Central Andes employing an ensemble machine learning model capable of accounting for interactions among predictors to assess whether the observed variation between individuals' diets results more from changing local climates or from population. Our findings suggest that climate change is the dominant driver of Central Andean dietary variation.

**Regional background.** The Central Andes are characterized by extreme heterogeneity in geography<sup>27,30</sup> and spatially and temporally diverse cultures<sup>31</sup>. Prehispanic subsistence ranged from marine and high-elevation foragers<sup>32–34</sup> to semi-sedentary pastoralists and agro-pastoralists<sup>35–37</sup> to semi- and mostly sedentary agriculturalists occupying multiple elevational niches<sup>38</sup>. Prior research has documented the potential for environment<sup>39</sup>, demography<sup>16,40</sup>, climate<sup>41</sup>, and sociocultural factors<sup>42,43</sup> to influence sociopolitical, economic, and dietary changes.

Climatically, the Central Andes experienced significantly changing conditions over the Holocene<sup>44</sup>. Variations in El Niño Southern Oscillation (ENSO) frequency<sup>45,46</sup> and resultant effects such as sediment deposition<sup>47</sup> and precipitation and drought<sup>48,49</sup>, have all heavily impacted the area. Spatial variation in the timing and intensity of ENSO appears to have existed between north and south Peru<sup>46</sup>, with staggered transitions in both regions from low intensity and dampened ENSO frequencies to modern ENSO frequencies over the past ~ 7000 years. Largely independent of ENSO, average temperatures in the Central Andes have also undergone changes and significant variation<sup>44,50</sup>, with sub-regions and localities exhibiting at times divergent patterns. Climatic variation in the Central Andes has been linked to socioeconomic changes<sup>47,51,52</sup>, intensification in production or infrastructure<sup>41</sup>, settlement shifts<sup>53</sup>, and population history<sup>29,54</sup>.

Few demographic reconstructions of population for the Central Andes exist, but those that do suggest populations were relatively low and stable early in the settlement of the region, with rapid increases in size later in time, though the rates and scales of population change differ by elevation<sup>29,40,55</sup>. Rapid population expansion<sup>56</sup> and larger population sizes<sup>57</sup> in the region appear to correlate with the incorporation of domesticates and transition to reliance on intensive agriculture as well as decreasing climatic volatility<sup>29</sup>. Here we generate a proxy of the past population sizes for each individual as a quantitative measure of the relative changes over time.

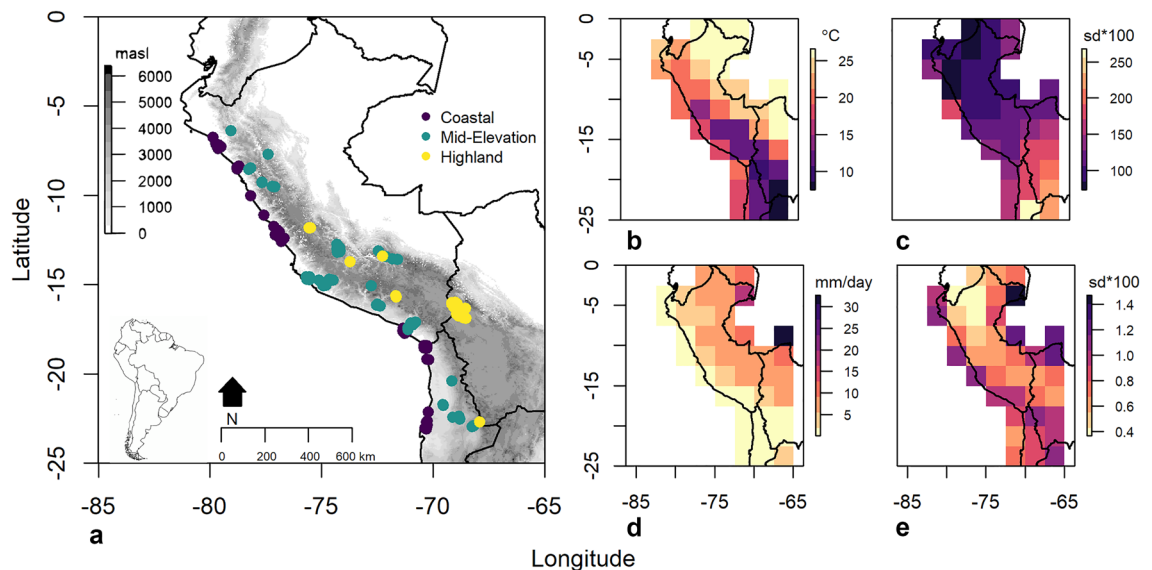
The sociopolitical trajectory, in general, of the Central Andes from the terminal Pleistocene until the arrival of the Spanish is one of increasing political interconnectivity. Particularly after ~ 4000 yBP, periods of greater and lesser political centralization occur, culminating in the region-wide Inca Empire (~ 480–418 yBP)<sup>31,58,59</sup>. Initial domestication of resources began before 6000 yBP<sup>38,60</sup>, predating the emergence of more complex sociopolitical entities and integration which began by 3000 yBP<sup>61,62</sup>. Socio-cultural practices around ritual/religion, sociopolitical complexity, trade, elite manipulation, and cultural continuity and hegemony have been well documented in the Central Andes, with significant differences in practices, beliefs, and effects over space and time<sup>63–66</sup>. The changes in sociopolitical complexity in particular may be linked with population dynamics, as increasing population size may increase sociopolitical complexity and drive the rise of the state<sup>16</sup> or vice-versa. Increasing sociopolitical complexity in the Andes produced varying patterns of influence and manipulation of resources and people by polities, including resettlement of individuals<sup>67,68</sup>, interconnectivity through trade<sup>69,70</sup>, and altering social dimensions of food<sup>43,71</sup>, all of which could influence individual diets.

## Results

Examining individual diets represented by  $\delta^{15}\text{N}\text{‰}$  or  $\delta^{13}\text{C}\text{‰}$  measured on bone collagen as a function of paleoclimate reconstructions from a GCM (Figs. 1, 2) and paleo-demographic reconstructions using Kernel Density Estimates (KDEs) (Fig. 2) depicts how the long-term dietary patterns in the Central Andes are influenced by climate and population.

To control for the elevation induced effects on climate<sup>27</sup> and subsistence strategies<sup>28</sup>, here individual diets and broader dietary change are analyzed across three elevation zones: coastal, mid-elevations, and highlands (see methods for details on assignment of individuals). These zones approximate elevation-influenced subsistence patterns<sup>28,31,35</sup> and allow for: a) reconstruction of general dietary change within similar strategies, and b) evaluation of if the effects of climate and population vary by subsistence strategy.

**Visualizing temporal variation in Central Andean diets by elevation zone.** To visualize temporal variation in diet, we collapse spatial variation into the three elevation categories and fit the trend using bootstrapped generalized additive model (GAM) regression. This reveals several patterns in broad dietary change over time across elevation zones (Fig. 3). (1) Unsurprisingly, diets differed significantly depending upon the elevation zone in which an individual lived. Until the Late Horizon period, coastal individuals generally have  $\delta^{15}\text{N}$  values 5 to 10‰ higher than all others. This is combined with 2 to 6‰ higher  $\delta^{13}\text{C}$  until about 1300 yBP. Mid-elevation individuals mostly have diets around 10‰  $\delta^{15}\text{N}$  over the entire period, though the typical  $\delta^{13}\text{C}$  values change drastically between ~ 2500 yBP and 1350 yBP. Highland individuals have similar  $\delta^{15}\text{N}$  values as those from mid-elevations but generally have  $\delta^{13}\text{C}$  values that are 2 to 4‰ lower than mid-elevation individuals



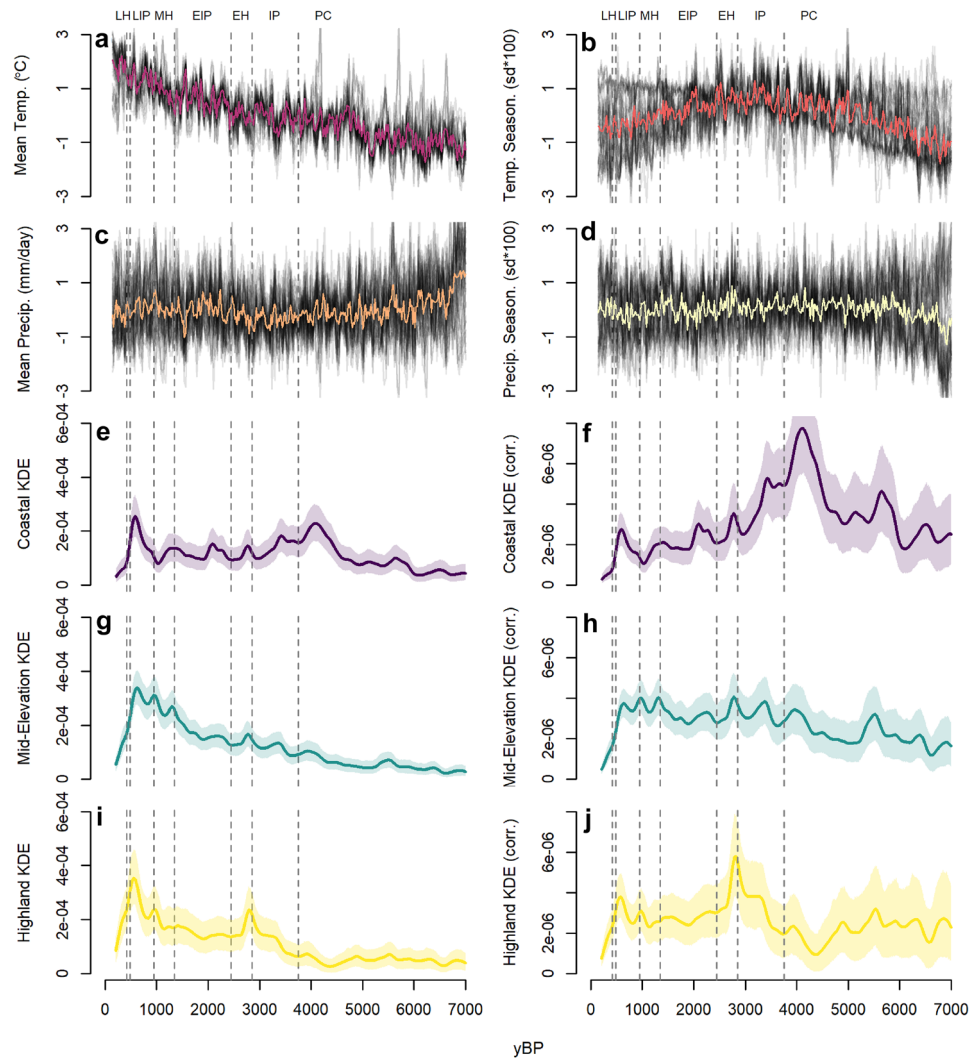
**Figure 1.** Study area and spatial climate variation. Left: (a) Map of locations from which individuals are derived. Colors correspond to the elevation zone. Right: Averaged climate value for each grid cell from the past 7000 years for the study area. Climate variables are (b) temperature ( $^{\circ}\text{C}$ ), (c) temperature seasonality (standard deviation in  $^{\circ}\text{C} * 100$ ), (d) precipitation (mm/day), and (e) precipitation seasonality (standard deviation in mm/day \* 100) generated from TraCE21ka simulations<sup>119,120</sup> through the PaleoView Paleoenvironmental Reconstruction Tool<sup>122</sup>. Maps were produced in the R statistical environment<sup>136</sup> using the **mapdata**<sup>150</sup> and **raster**<sup>151</sup> packages with shapefiles obtained from the NOAA Global self-consistent, hierarchical, high-resolution geography database<sup>152</sup>.

and up to 6‰ lower than those on the coast. (2) An increase in  $\delta^{13}\text{C}$  values over time is noticeable across each elevation zone, with the largest change in the mid-elevations. This increase is smallest on the coasts and likely occurs sometime between the Preceramic and Early Horizon. In the mid-elevations and highlands, the increase occurs later in time, appearing to happen during the Early Intermediate Period, with mid-elevation values generally experiencing a ~6‰ increase with highland values increasing by about 3‰. (3) The central tendency in  $\delta^{15}\text{N}$  for mid-elevation individuals appears to possess some variation dependent upon period, in particular showing elevated values 2 to 4‰ higher during the Early and Late Intermediate Periods relative to the Middle and Late Horizons. (4) The amount of differentiation between individuals' diets diminishes during the Late Horizon. During this period,  $\delta^{15}\text{N}$  values across each elevation overlap with each other around 10 to 11‰. At the same time,  $\delta^{13}\text{C}$  values for mid-elevation and coastal individuals are quite similar, around -11 to -12‰, though highland values continue to remain 3 to 4‰ lower. These patterns are also visible in 95% confidence interval ellipses plots (Supplementary Fig. 1, Supplementary Material 2).

**Climate and population predict past diets.** The cross-validated random forest (RF) regression models evaluating the predictive power of climatic and population change on diet reveal that combined, our climate and population variables explain 79.82 and 66.02% of the deviance in  $\delta^{15}\text{N}\text{‰}$  and  $\delta^{13}\text{C}\text{‰}$  respectively when using uncorrected KDEs, or 80.04% and 66.77% when correcting for taphonomic loss. Root mean square error values for  $\delta^{15}\text{N}\text{‰}$  models are 2.05 (uncorrected KDE) or 2.04 (corrected KDE) and are 1.62 or 1.60 for  $\delta^{13}\text{C}\text{‰}$ . As such, combined climate and population estimates explain the majority of dietary variation, regardless of demographic estimate employed. Model diagnostics reveal residuals are normally distributed (Supplementary Figs. S6, S8, S10, S12) around zero, possess limited, unstructured temporal autocorrelation to about 200 years (Supplementary Figs. S7, S9, S11, S13), and do not possess spatial autocorrelation, indicating that the models perform well in predicting the variation in  $\delta^{15}\text{N}\text{‰}$  and  $\delta^{13}\text{C}\text{‰}$  across time and space.

**Climate has a stronger influence on diets than population.** As the models perform well in predicting diet, we employ effect size calculations for each predictor variable within each elevation zone to evaluate whether climate or population change has a greater impact. Effect sizes, controlled for the interactions of all predictor variables, reveal climate has a much larger effect on  $\delta^{15}\text{N}\text{‰}$  and  $\delta^{13}\text{C}\text{‰}$  than population (Tables 1, 2, Fig. 4). Climate, cumulatively, has an effect size ranging from two to fourteen times that of demography across both isotopes.

Regardless of demographic proxy employed (taphonomically uncorrected or corrected KDE), climate explains the majority of the variation in the isotopic values between individuals across space and time. Even considering climatic variables independently suggests that individual climatic factors are more influential than demography in every elevation zone for both isotopes (Tables 1, 2). The one analytical location where the influence of demography comes close to that of climate is mid-elevation  $\delta^{15}\text{N}\text{‰}$ , where, across the temporal range of observations,



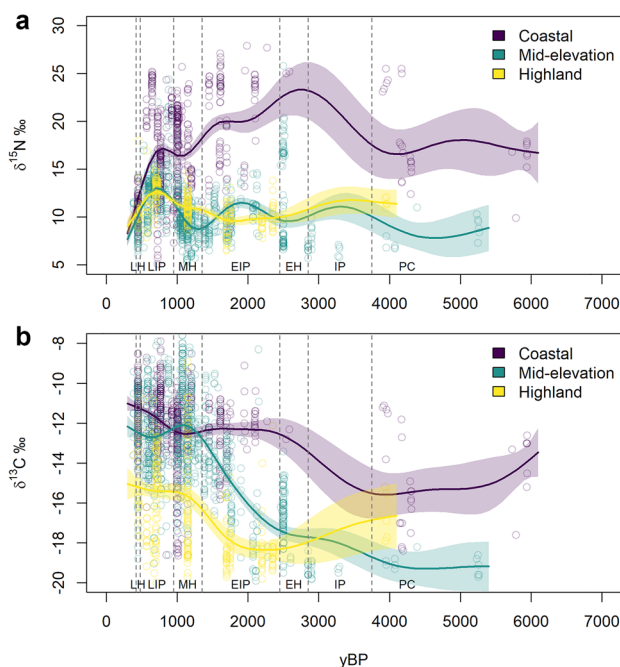
**Figure 2.** Time series of climate and demographic variables. (a–d) z-scores of temporal climate deviation from the mean for the past 7000 years. Each grey line represents the z-score variation of an individual grid cell (see Fig. 1). Colored lines represent the moving average z-score for all raster grid cells combined. In order, the variables are (a) mean temperature ( $^{\circ}\text{C}$ ), (b) temperature seasonality ( $\text{sd } ^{\circ}\text{C} \times 100$ ), (c) mean precipitation ( $\text{mm/day}$ ), and (d) precipitation seasonality ( $\text{sd mm/day} \times 100$ ). (e–j) Coastal, mid-elevation, and highland KDE estimates documenting relative variation in population size over time for each elevation category employing uncorrected (left) and taphonomically corrected (right) KDEs. For analytical purposes, individuals receive climate value estimates from the time series of the grid cell at their spatial coordinates, not the moving average central tendency line.

the cumulative effect of climate is a little over two times (Table 1) that of demography, measured as uncorrected KDEs, and the effect size of demography is  $\sim 3.8\%$ .

Individually, precipitation is the most influential climate variable for coastal and mid-elevation  $\delta^{15}\text{N}\%$  regardless of KDE employed, whereas in the highlands, precipitation seasonality has the largest effect on  $\delta^{15}\text{N}\%$  (Tables 1, 2). For  $\delta^{13}\text{C}\%$ , each of the climate variables tend to be similarly correlated in the coastal and mid-elevations, with mean temperature having a larger relative effect in the highlands (Tables 1, 2).

When using uncorrected KDEs, population size fairly consistently correlates with around 13% of the variability, across isotope and elevation category, with the noted exception of mid-elevation  $\delta^{15}\text{N}\%$  where demography corresponds with 22.1% ( $\sim 3.8\%$ ) of the variability. When using corrected KDEs, population size generally accounts for around 10% of the variability, maxing out at 11.61% (0.60‰) for highland  $\delta^{15}\text{N}\%$ .

In terms of per mil variability, the demographic effect corresponds with potentially meaningful variation ( $> 1.0\%$ ) in mid-elevation  $\delta^{15}\text{N}$  (3.8‰) and  $\delta^{13}\text{C}$  (2.7‰) as well as coastal  $\delta^{15}\text{N}$  (1.6‰) when using uncorrected KDEs. When using corrected KDEs, the most meaningful effects of demography are on coastal and mid-elevation  $\delta^{15}\text{N}$  (1.4 and 1.3‰ respectively). However, these values are smaller than most of the individual climate variable effects and much smaller than the respective cumulative effects of climate. See Supplementary Tables S1 to S4



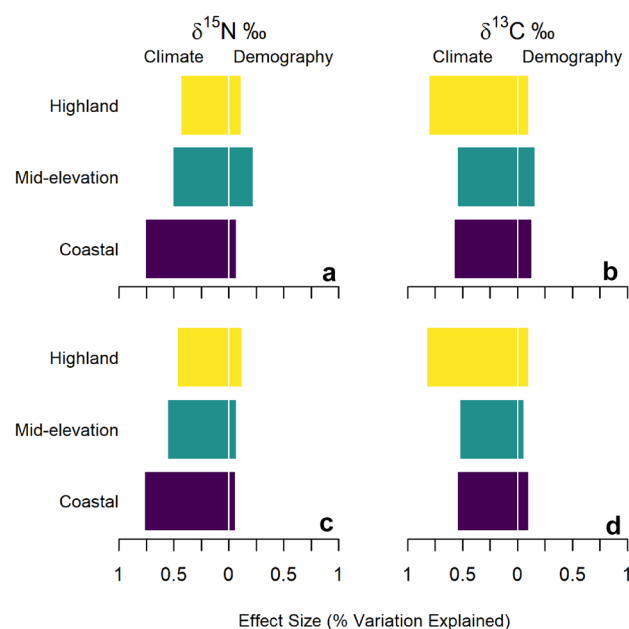
**Figure 3.** Generalized additive model time series. GAM regression results for both (a)  $\delta^{15}\text{N}\text{‰}$  and (b)  $\delta^{13}\text{C}\text{‰}$  across the three elevation categories. Dots represent the observed individuals plotted at their median date determined via resampling of their date range 10,000 times (weighted by calibrated radiocarbon probability for directly dated individuals). Solid lines are the central tendencies for each elevation category with shaded 95% confidence intervals. Lines were generated by creating 10,000 GAMs per elevation zone and isotope (60,000 total) where, for each GAM, each individual received a single year as their date, sampled via weighted sampling of their date range. The central tendency line is the mean fit line of the 10,000 GAMs for each sample, with 95% CIs created from the mean standard error and standard deviation of the 10,000 GAMs for each sample. Time periods follow the chronological periods from Moseley<sup>58</sup>: Preceramic (PC), Initial Period (IP), Early Horizon (EH), Early Intermediate Period (EIP), Middle Horizon (MH), Late Intermediate Period (LIP), Late Horizon (LH). Note: This is a time series only and does not incorporate space. Some of the variation within elevation zones relates to spatial variation which is addressed in the RF analyses.

Isotope	Elevation category	Temp. EF‰ (%)	Temp. seasonality EF‰ (%)	Precip. EF‰ (%)	Precip. seasonality EF‰ (%)	Cumulative climate EF‰ (%)	Demography EF‰ (%)
$\delta^{15}\text{N}\text{‰}$	Coastal	1.4 (5.6%)	6.5 (25.7%)	10.2 (40.0%)	1.2 (4.8%)	19.4 (76.2%)	1.6 (6.6%)
	Mid-elevation	0.7 (4.1%)	2.0 (11.8%)	5.3 (29.9%)	0.9 (5.5%)	9.0 (51.2%)	3.8 (22.1%)
	Highland	0.4 (8.5%)	0.4 (9.3%)	0.5 (11.0%)	0.7 (15.4%)	2.1 (44.1%)	0.5 (11.2%)
$\delta^{13}\text{C}\text{‰}$	Coastal	0.7 (10.3%)	0.7 (10.2%)	1.5 (21.0%)	1.2 (16.7%)	4.1 (58.1%)	0.9 (13.3%)
	Mid-elevation	1.0 (6.4%)	2.6 (15.9%)	1.8 (11.0%)	3.5 (21.2%)	8.9 (54.5%)	2.7 (16.4%)
	Highland	2.2 (35.4%)	0.6 (9.6%)	1.1 (18.7%)	1.1 (17.1%)	4.9 (80.7%)	0.6 (10.4%)

**Table 1.** Random Forest Model Effect Sizes with uncorrected KDEs. Model sum of squares effect sizes (amount of variation in  $\delta^{15}\text{N}\text{‰}$  or  $\delta^{13}\text{C}\text{‰}$  explained) for each variable within the elevation categories. These values are in per mil (‰). Percentages are the model sum of squares effect sizes as a percent of the total amount of variation. Effect sizes are directly comparable within each isotope, but not between them, given that there are different models for each stable isotope. Values here have been deducted by Friedman’s *H*-statistic to remove the proportion of the effects resulting from interaction between variables. Here demography is generated via uncorrected KDEs. Note: the effect sizes are calculated through iterated simulation employing representative subsets of each key predictor variable, as they do not include every possible combination the percentages will not sum to 100.

Isotope	Elevation category	Temp. EF‰ (%)	Temp. seasonality EF‰ (%)	Precip. EF‰ (%)	Precip. seasonality EF‰ (%)	Cumulative climate EF‰ (%)	Demography EF‰ (%)
$\delta^{15}\text{N}\text{‰}$	Coastal	1.7 (6.7%)	6.7 (26.4%)	10.1 (39.4%)	1.3 (5.0%)	19.7 (77.4%)	1.4 (5.7%)
	Mid-elevation	1.2 (6.9%)	2.0 (11.5%)	5.0 (28.7%)	1.5 (8.8%)	9.7 (55.9%)	1.3 (7.4%)
	Highland	0.5 (10.8%)	0.6 (12.5%)	0.4 (9.4%)	0.7 (14.2%)	2.2 (47.0%)	0.6 (11.6%)
$\delta^{13}\text{C}\text{‰}$	Coastal	0.8 (10.9%)	0.7 (10.6%)	1.2 (17.7%)	1.1 (15.9%)	3.9 (55.1%)	0.7 (10.1%)
	Mid-elevation	1.0 (6.5%)	2.8 (17.0%)	1.5 (9.0%)	3.3 (20.3%)	8.6 (52.7%)	1.0 (5.9%)
	Highland	2.1 (34.9%)	0.6 (10.6%)	1.1 (18.8%)	1.1 (18.4%)	5.0 (82.7%)	0.6 (10.4%)

**Table 2.** Random forest model effect sizes with corrected KDEs. Model sum of squares effect sizes (amount of variation in  $\delta^{15}\text{N}\text{‰}$  or  $\delta^{13}\text{C}\text{‰}$  explained) for each variable within the elevation categories. These values are in per mil (‰). Percentages are the model sum of squares effect sizes as a percent of the total amount of variation. Effect sizes are directly comparable within each isotope, but not between them, given that there are different models for each stable isotope. Values here have been deducted by Friedman's *H*-statistic to remove the proportion of the effects resulting from interaction between variables. Here demography is generated via taphonomically corrected KDEs. Note: the effect sizes are calculated through iterated simulation employing representative subsets of each key predictor variable, as they do not include every possible combination the percentages will not sum to 100.



**Figure 4.** Cumulative climate and demographic effect sizes. Cumulative climate and demographic effect sizes (percent of the per mil amount of variation explained) within each elevation category for  $\delta^{15}\text{N}\text{‰}$  (left column) and  $\delta^{13}\text{C}\text{‰}$  (right column). (a, b) The cumulative effect sizes employing the uncorrected KDEs. (c, d) The cumulative effect sizes employing the taphonomically corrected KDEs. Regardless of KDE used for demographic estimates, climate change consistently has a larger effect than demography.

for the complete set of Friedman's *H* statistics and both initial and deducted effect size values for each variable within each elevation category.

## Discussion

Reconstruction of the long-term dietary trends in the Central Andes reveals significant dietary variation existed, particularly between, though also within, elevation zones. Given the isotopic differences between marine, terrestrial mammal, and terrestrial plant resources (see methods for overview), variation between elevation zones is enhanced by what seems to be multi-trophic level difference from coastal to all other individuals. Based on  $\delta^{15}\text{N}\text{‰}$ , this difference is likely driven by differential marine resource consumption. Ready access to and reliance on marine resources is well-documented<sup>72–75</sup>, and should increase relative  $\delta^{15}\text{N}\text{‰}$ . Until around 1350 yBP,  $\delta^{13}\text{C}\text{‰}$  shows significant differences between coastal and other individuals as well. While marine resources certainly contribute to this pattern given their typical values<sup>76,77</sup>, the  $\delta^{13}\text{C}\text{‰}$  differences may also suggest different levels of reliance on C3 vs C4 plants, with mid-elevation individuals more heavily C3 reliant up until around 1350 yBP.

By ~ 1350 yBP, increasing mid-elevation and highland  $\delta^{13}\text{C}\text{‰}$  may suggest an increase in maize consumption (as food or as *chicha*), which has been suggested to have taken longer to establish in the mid- and high elevations and southern latitudes than on the more northerly coasts<sup>38,78,79</sup>. Alternately, it could be evidence of greater inclusion of low-trophic level marine resources farther inland<sup>70</sup> or of animal husbandry practices employing greater amounts of C4 fodder<sup>80</sup>. Consistently however, highland individuals possess lower  $\delta^{13}\text{C}\text{‰}$  values than all others, suggestive of continual use of more C3 resources than in the other zones. This may be expected, particularly as the ceiling for maize cultivation today is generally ~ 3400 masl, though specific microenvironments enable cultivation up to ~ 3800 masl<sup>7,38,81</sup>.

Our results suggest the observed differences in individuals' diets are driven by local climatic conditions: climate has a much stronger correlation than relative population size with stable isotope measured individual diets in the Central Andes. Crucially, for the analysis, our results are derived from local spatiotemporal measures. This means that while the central tendency in dietary change (Fig. 3) captures the general temporal trends, our results (Fig. 4) reveal the predictive power of climate and population when considering spatial as well as temporal variation. The pattern of climate having more influence than population holds whether considering climate variables individually or their cumulative effect and regardless of whether corrected or uncorrected demographic estimates are employed (Fig. 4, Tables 1, 2). Given that the RF models explain ~ 80% of the variance in  $\delta^{15}\text{N}\text{‰}$ , and ~ 66% in  $\delta^{13}\text{C}\text{‰}$ , and that climate is the significantly stronger correlate, our results suggest that local climate strongly constrains viable subsistence adaptations in the Central Andes. As the local climatic conditions experienced by individuals changed through time, individuals adapted by modifying their diets or subsistence strategies. Given our inability to control for foddering or fertilization for each of our individuals, and our reliance exclusively on  $\delta^{13}\text{C}\text{‰}$  and  $\delta^{15}\text{N}\text{‰}$ , we cannot yet determine if this correlation is due to changing food items in diets or changes in subsistence practices that maintain the same food items but alter their isotopic signature. Likely both occur and both may produce large changes in the isotopic signature of resources. For instance, some of the dietary changes over time may represent greater incorporation of new resources, such as maize<sup>60,82–84</sup>. However other dietary changes may result from the introduction of new practices like fertilization or increased animal foddering<sup>80,85–87</sup>. If these new subsistence practices were responses to climate change, as has been suggested for multiple changes in practices<sup>41,47,51,52</sup>, this would be captured in our analysis as climatic effect on diet even though the foods being consumed did not change.

The significantly stronger impact of climate than population on isotope value, and the fact that the majority of our samples come from the past 2000 years, may suggest that vertical migration and horizontal exchange between groups in the Central Andes<sup>54,88</sup>, intensification in response to heightened levels of violence and warfare<sup>89</sup>, and innovations in subsistence economy techniques<sup>41,51</sup> relate to local climate more strongly than demographically induced competition or sociopolitical change. However, demography, regardless of corrected or uncorrected population proxy, does consistently account for around 10–13% of the effect, indicating population changes exert some influence. Our results suggest this effect may be largest for mid-elevation  $\delta^{15}\text{N}\text{‰}$ . A larger mid-elevation effect may relate to the heavier dependence on agriculture as a key subsistence strategy as competition over arable land may lead to violence<sup>89,90</sup>, state level formation<sup>30</sup>, co-option of land<sup>91</sup>, or extensification; all of which could interact with population size to have a bigger influence on mid-elevation diets. Further, it is plausible that the incised drainages of the mid-elevations allowed changes in population size to have more dramatic influences on population density than in other areas. Intensification in response to density changes that produces greater reliance on agriculture through fertilization<sup>92</sup> or terrace building<sup>93</sup>, as well as trade<sup>69</sup>, or other behaviors such as water engineering may be more necessary for Central Andean mid-elevation agriculturalists than for those individuals with access to more marine or pastoral resources. It is also plausible that demographically driven competition caused intensification when and where climatic constraints allowed, such as in the middle elevations, but caused scarcity in others. Future work evaluating if this 10%, or larger in the case of mid-elevation  $\delta^{15}\text{N}\text{‰}$ , effect is consistent over time and space or driven by stronger population influences during a particular period(s) may unveil new insights into the temporal dynamics of population influence in the Andes. For instance, it is plausible that alteration in subsistence techniques could occur coincident with climate change, but not in response to it (i.e., Boserupian intensification resultant from increasing population separated from climatic effects), and future studies may be able to elucidate if this occurs at various points in time in the Central Andes.

Divergent from the rest of the time periods however, during the Late Horizon isotope values from most mid-elevation and coastal, as well as some highland, individuals overlap. Individuals living on the coast experienced a sharp drop in  $\delta^{15}\text{N}\text{‰}$ , with a similar albeit smaller decrease experienced by mid-elevation and highland individuals. Particularly on the coast this is a surprise as it appears to be a decrease in either the amount or trophic level of marine resources consumed, despite a near 7000-year history of coastal individuals making heavy use of high  $\delta^{15}\text{N}\text{‰}$  marine resources. At the same time, mid-elevation and coastal  $\delta^{13}\text{C}\text{‰}$  continued to be nearly indistinguishable. This stark decrease in dietary variation between individuals occurs temporally rapidly over broad elevation gradients. These findings may result from a warming less seasonal environment, which could plausibly decrease variability in resources while increasing the economic payoffs from agriculture, or from sociopolitical change.

Overall, earlier in time, climate appears to have been more locally disparate (Fig. 2), with somewhat greater deviations from the mean in temperature, temperature seasonality, mean precipitation, and precipitation seasonality than later in time. In the context of marked spatial variation in climates, local resources would have been varied in type, presence and abundance. Over the past 1300 years, however, temperatures in the region continued to rise significantly above the mean, temperature seasonality generally decreased from the preceding 1000 years, and precipitation and precipitation seasonality became slightly more spatially varied from preceding periods, with some locations experiencing decreased precipitation and seasonality while others saw increases (Fig. 2). Taken together, a warmer and potentially less seasonal environment may have decreased some of the variability in resources, either latitudinally or longitudinally, while increasing the viability of agriculture as a

subsistence strategy, which could cause the differences between individuals' diets to diminish. Diminishing differentiation was potentially occurring between ~ 1300 and 500 yBP and appears to have strongly occurred during the Late Horizon.

Alternatively however, dietary convergence may support earlier work by Burger, et al.<sup>94</sup> suggesting Inca administration may have led to heavy reliance on maize across the Central Andes, though their conclusions were based dominantly on individuals from Machu Picchu. Given the spatial spread and political influence of the Inca, such convergence in the isotope data likely reflects the power of the Inca Empire, realized through state level redistribution of resources<sup>69,95,96</sup> and people<sup>67</sup>, as well as emphasis on the social importance of maize<sup>71</sup>. These activities may have collapsed dietary variation across elevation zones and may reflect increasing homogenization in subsistence practices. The overall pattern of dietary convergence coinciding with increasing sociopolitical integration and potential genetic homogenization<sup>97</sup> suggests the possibility of strong interplay between social processes and consumption late in time.

During the Middle Horizon (~ 1350–950 yBP), both the Tiwanaku and Wari Empires engaged in regional integration, trade, and resettlement<sup>98,99</sup> as well, and they contributed to changing social dimensions of maize<sup>43,65</sup>, yet individuals' diets remained quite distinct both between and within elevation zones during this period. Mid-elevation and coastal  $\delta^{13}\text{C}\text{‰}$  does overlap, but  $\delta^{15}\text{N}\text{‰}$  stays about 5 to 8‰ higher on the coasts even as inland transport of marine resources was possible and occurring, at least in southern Peru<sup>70</sup>. Additionally, during the Late Intermediate Period (~ 950–480 yBP), vertical migration between groups<sup>54</sup> and horizontal exchange<sup>88</sup> suggest economic and/or political as well as social interconnectivity between individuals in different elevation zones existed. This may be expected to significantly decrease dietary variation and, yet, large dietary differentiation between and within elevation zones remained. Thus, our results may suggest that either integration between the coast and mid-elevations was not widespread or may have been limited in its scope during the Middle Horizon and Late Intermediate Period. Taken together with the pattern during the Late Horizon, our results may suggest that, in the Late Horizon, the influence exerted by the Inca Empire overrode local climatic influences on diet in ways the Wari and Tiwanaku Empires and the increased Balkanization of the LIP could not. Alternately, given the strength of climate in our analysis, it is possible that continued climate change drove optimal socioeconomic decisions causing subsistence behaviors to fixate on similar resources. Finally, there is also potential for unique interactions between climate change and sociopolitical processes in the Late Horizon.

The sharp convergence of diets during the Late Horizon and the ~ 20–35% of the variation remaining unexplained by our models suggests that though climate change is the dominant influence it is not the sole explanatory factor. The remaining variation may be representative, in part, of the effect that sociopolitical complexity and other social factors had on dietary variation. There are several possible explanations for how sociopolitical complexity could account for remaining variation even though demographic influence within the models is small. First, it may be that the effects of sociopolitical complexity on dietary variation are divorced from population effects, meaning the number of people involved is less important a factor than the influences exerted by state-level decisions, organizational complexity, changing ritual practices, market integrations, etc., particularly if those factors are not driven by population change. This could be especially true if sociopolitical institutions were adapting to climate change by promoting behaviors such as exchange, redistribution, or increased forays into multiple elevation zones. Such adaptations in response to climate change would result in decreasing dietary variation that correlates strongly with climate but not population size. Second, it is possible that population does have a larger effect and correlate with sociopolitical complexity but that our climate proxies more accurately estimate past climate than our demographic proxies estimate relative past population sizes. Finally, it is possible, maybe likely, that our population reconstructions underestimate relative population size during the Late Horizon in particular, which could decrease the influence of demography if it were having a significantly stronger impact on individuals during that period than over the rest of the study period. Each of these possibilities warrants future investigation.

To further evaluate whether Central Andean dietary change may be driven by changing foods or subsistence practices and what may be causing the Late Horizon dietary overlap, significant future work is required. Estimates of sociopolitical influence at the site level both within and between temporal periods along with increased sample sizes of multiple different light isotopes, greater temporal control for most individuals within the region, improved climate change proxies, expanded datasets on non-isotope proxies for diet (i.e., zooarchaeology, paleobotany), and increased multi-site investigations will all be necessary. Combining these various datatypes will prove useful in generating unique insights into the long-term human–environment interactions in the Central Andes.

Here we show that in a region defined by its unique cultural, elevation, subsistence, and climatic heterogeneity, the local climatic conditions individuals experienced during their lives were the strongest factors influencing their dietary patterns. Our models employing simulated climate with relative population size explain ~ 80% of the variation in  $\delta^{15}\text{N}\text{‰}$  and ~ 66% in  $\delta^{13}\text{C}\text{‰}$  across the past 7000 years. While we anticipate more detailed reconstructions of local climate change, more nuanced demographic reconstructions, and future incorporation of social factors may increase the amount of the dietary variation explained, these current variables perform well in constraining isotopic values of diet. Our results suggest that climatic changes are a stronger correlate with dietary differentiation between individuals than changing population sizes. Though we suspect sociopolitical complexity impacted dietary variation as well, its influence appears constrained within a pattern of climatic change and environmental limitation, at least until the Late Horizon. Such a result implies that past and future climatic change did and will highly influence subsistence decisions and dietary outcomes. Given the rapid climate changes occurring in the world today, our analysis of dietary change over the past 7000 years in the Central Andes implies that generating projected climatic changes will be highly productive in predicting health<sup>17</sup> and complexity<sup>19</sup> changes in the future as these aspects of life are intimately impacted by subsistence.



The approach presented here offers a systematic framework for exploring the relative role of climate and other socio-demographic factors on dietary change through time. We suspect that climate will have a similarly outsized influence on long-term dietary change in other regions where systematic studies have yet to be undertaken.

## Materials and methods

**Data.** We compiled a database of 1965 individuals (Supplementary Data File 1) from published literature in Peruvian, northern Chilean, and Lake Titicaca archaeological contexts (Fig. 1) with  $\delta^{13}\text{C}\text{‰}$  or  $\delta^{15}\text{N}\text{‰}$  measured on bone collagen. Nearly 40 years of isotopic studies have documented that carbon and nitrogen isotope ratios from human bone collagen in the Central Andes are responsive to the proportion of C3/C4 plants, marine/terrestrial animal input, and other dietary behaviors, effectively capturing variation in diets<sup>100–104</sup>. For our analysis, we exclude individuals under the age of five to control for weaning effects but include other subadults noting that although subadult isotope data represent a shorter window of life compared to adults, these data are independent of nursing induced trophic level enrichment<sup>105,106</sup>. Individuals are further evaluated using reported C:N ratios; those with C:N ratios between 2.9 and 3.6<sup>107</sup> are accepted as reliable data. Each individual outside that ratio is checked for the reporting author's decision on whether to accept the isotope values as valid. If the reporting author accepts a value outside the 2.9–3.6 ratio we include that individual in the dataset. In effect, this process expands the acceptable C:N ratio to 2.6–3.6, as many authors rely upon the 2.6–3.4 C:N ratio proposed by Schoeninger, et al<sup>108</sup>. Some individuals do not have a direct C:N ratio, but the reporting author confirmed in text that all individuals were within the acceptable C:N range, resulting in inclusion in our analysis. If authors did not report a C:N ratio, specifically or study-wide, but accept their data as valid, we include the data ( $n = 122$ , 7% of the sample). As a final step, we drop Colonial Period individuals ( $n = 25$ ) to constrain the analysis to pre-Spanish arrival. This process results in a sample of 1767 individuals, 1727 with bone collagen  $\delta^{15}\text{N}$  values and 1761 with bone collagen  $\delta^{13}\text{C}$  (Supplementary Data File 2). These isotope values, which represent each individuals' average adult diet<sup>105</sup>, span the past ~7000 years.

Elevation in the Andes affects climate<sup>27</sup> and subsistence modes/opportunities<sup>28</sup>. Therefore, we group individuals into three elevation categories (coastal, mid-elevation, and highland) which roughly correspond to elevation-influenced subsistence patterns: agro-marine with heavier marine influence on the coast, agro-marine-pastoral in the mid-elevations with greater emphasis on agriculture, and agro-pastoral with heavier influence of pastoralism in the highlands<sup>28,31,35</sup>. This allows us to i) evaluate the long-term trends in dietary variation across different subsistence strategies, ii) more directly examine whether climate or population have greater effects on dietary variation by examining influences within elevation categories, and iii) evaluate if the effects of climate and population size vary between the categories, implying differential influences based on primary subsistence pattern.

We establish the elevation category (coastal, mid-elevation, highland) for each individual by extracting elevation (meters above sea level, masl) from a digital elevation map (DEM) and employing an adapted version of the natural zones from Pulgar Vidal<sup>27</sup>. Any individual recovered from less than 350 masl and a distance to coast of less than 15 km is categorized as coastal. Individuals recovered from less than 3500 masl, but above 350 masl, or less than 350 masl but greater than 15 km from the coast are considered mid-elevation. Individuals recovered from above 3500 masl are categorized as highland. Defining elevation zones places the measure of elevation on a level of precision comparable to our climatic and demographic data and allows us to control for some of the spatial variation to enhance analyzing variation over time.

**Stable isotopes.** Isotope values of bone accurately record isotopic composition of dietary inputs with known fractionation offsets and thus can be used to examine dietary variation between groups and individuals<sup>109–111</sup>. Here we rely solely on bone collagen  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  per mil values as our inferred dietary measure to make all individuals directly comparable and maximize sample size. Based primarily on the differences in photosynthetic discrimination against metabolism of  $^{13}\text{CO}_2$  in C3 versus C4 plants<sup>109,112</sup>, with limited variation due to local soil chemistry and aridity<sup>113</sup>,  $\delta^{13}\text{C}$  values track reliance on C3 versus C4 plant foods and animal protein sources reliant on such foods. In secondary consumers like humans,  $\delta^{15}\text{N}$  values are useful for revealing trophic level and protein intake<sup>110,114,115</sup>. With each step up a food web,  $\delta^{15}\text{N}$  fractionation increases by ~2–4‰<sup>115</sup>. Though overlapping  $\delta^{13}\text{C}$  values may exist between marine mammals, terrestrial plants, and terrestrial mammals, the combination of  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  typically enables discrimination between dietary sources<sup>116</sup>.

In the Central Andes, distinct isotopic signatures of key resources have proven useful in identifying variation in dietary reliance on maize (*Zea mays*) or other C4 plants<sup>60,82,103</sup>, marine mammals and fish<sup>76</sup>, camelids<sup>104</sup>, and potatoes (*Solanum sp.*) and quinoa (*Chenopodium quinoa*)<sup>117</sup> among other resources. In general, maize in the Central Andes has  $\delta^{13}\text{C}$  signatures of ~–11 to –12‰ and  $\delta^{15}\text{N}$  of ~4 to 8‰<sup>85,117</sup>. The high  $\delta^{13}\text{C}$  values enable maize to be distinguished from C3 plants which have values such as ~–26‰ for potatoes, ~–25‰ for quinoa, and ~–25‰ for beans (*Phaseolus sp.*)<sup>85,117</sup>. Similarly, the isotopic signatures of Central Andean camelids differ from those of marine resources in that marine resources possess both elevated  $\delta^{13}\text{C}$  (~–11 to –16‰) and  $\delta^{15}\text{N}$  (~16 to 23‰) values<sup>76,77</sup>, whereas camelids tend to be less elevated in  $\delta^{15}\text{N}$  (~5 to 8‰), though  $\delta^{13}\text{C}$  may vary<sup>77,118</sup>. For this analysis, we do not explore the specific inputs contributing to observed  $\delta^{13}\text{C}\text{‰}$  and  $\delta^{15}\text{N}\text{‰}$  values. Rather, we use the observed values as composite measures of variation in diet.

**Climatic Variables.** We capture variation in local climates in space and time through a general circulation model simulation of mean annual precipitation (mm/day), mean annual temperature (°C), mean annual temperature seasonality (sd °C \* 100), and mean annual precipitation seasonality (sd mm/day \* 100) (Fig. 1). Pre-historic climatic variables are generated from the TRaCE-21 ka experiments<sup>119–121</sup> via the PaleoView Paleoenvironmental Reconstruction Tool<sup>122</sup>. Through this tool, we obtain estimates for precipitation, temperature, and

seasonality at 20-year intervals from 11,000 to 140 yBP over a  $2.5 \times 2.5$  degree latitude/longitude grid, providing broad resolution variation in spatial and temporal estimates (Figs. 1, 2).

Climate variables are related to each individual through iterated, weighted sampling of the climatic values within an individual's date range, using site coordinates to get the data from the individual's spatial location. Individual date ranges are based upon reporting author assignments, converted into years before present (yBP), or, for individuals directly radiocarbon dated ( $n = 155$ , ~9% of the sample), the calibrated radiocarbon date range BP. To help account for potential bias driven by uncertainty in the age estimate of each individual, we conduct a Monte Carlo sampling routine to obtain climate estimates. Each climate variable's values are extracted from each year within the maximum and minimum possible date for each individual. Each individual's temporal range is then randomly sampled, with replacement, 10,000 times. For individuals who are radiocarbon dated, the probability of selecting a given year's climate observation is weighted by the probability density of that year from the calibrated date. For individuals who are not radiocarbon dated, each year in their time window has equal probability of being selected at each sampling. We then average the 10,000 samples for each climatic variable per individual, providing an estimate of the most likely climatic conditions from their lifetime.

**Demography.** Evaluation of archaeological population size/demography is difficult. Here we attempt to capture relative changes in population size by proxy through recreating demographic trends as the relative change in population estimate between points in time. To evaluate this variation, we employ the “dates as data” approach<sup>29,40,123–125</sup>. To generate the data for this approach we collate radiocarbon dates from existing compilations generated by Ziolkowski<sup>126</sup>, Gayo, et al.<sup>55</sup>, Rademaker, et al.<sup>127</sup>, Goldberg, et al.<sup>128</sup>, Riris<sup>40</sup>, and Roscoe, et al.<sup>16</sup>. From the compilation, we select all dates from Peru, northern Chile, and the Bolivian highlands, to align with regions from which we have isotope data. Dates are checked for duplication using the laboratory ID code and all duplicated dates are reduced to one observation. We employ several strategies for date quality control. 1) We remove dates older than 15,000 years as unlikely to be human generated (though our analysis relies on dates from the past 7000 years and is not influenced by this removal). 2) Following Riris<sup>40</sup> and Robinson, et al.<sup>129</sup>, dates with errors greater than 200 years are also removed as such dates may be unreliable. 3) We drop any dates for which coordinates could not be identified. Dates from Roscoe, et al.<sup>16</sup> are an exception and are kept even if missing coordinates as the compilation relied exclusively on coastal dates. The quality control process results in a set of 3957 dates from which the demographic data are generated.

In order to account for uncertainty in both population size estimates derived from radiocarbon dates and uncertainty in the date in which each individual died, we employ composite kernel density estimates (KDEs) of radiocarbon date demographic reconstruction<sup>130–133</sup> and Monte Carlo iterated sampling. We generate population size estimates via KDEs within each of our elevation zones, allowing us to estimate population size trends separately for the coast, mid-elevations, and highlands. To classify dates by elevation, we take the coordinates for each date, extract elevation from the DEM, and then categorize the dates using the same elevation criterion as for individuals with isotope values. The resultant samples are 1304 ( $1115 < 7000$  yBP) coastal dates, 1843 ( $1610 < 7000$  yBP) mid-elevation dates, and 810 ( $657 < 7000$  yBP) highland dates from which we generate the KDEs. KDE generation requires calibrated dates; all non-marine dates are calibrated using SHCal20. Marine-derived dates are calibrated using the Marine20 curve with reservoir effect offsets for each date calculated using calib.org<sup>134</sup>. Though implementation of the SHCal20 curve for all dates may introduce some errors given the potential for mixing Northern and Southern hemisphere climate dynamics, the magnitude of any such errors will be small and distributed across our sample<sup>135</sup>.

Following previous researchers using dates as data in the Central Andes<sup>16,29,40</sup>, we employ hierarchical clustering using a defined cutoff of 200 years to control for overrepresentation of individual sites<sup>40,125,131</sup>. To construct KDEs, we employ the **rcarbon** package<sup>131</sup> in the R statistical environment<sup>136</sup> to generate 1000 unique kernel density estimates through random sampling of calendar ages from each of the calibrated radiocarbon probability distributions, providing an estimate of the envelope of possible reconstructions. Demographic estimates for each individual are then constructed through iterated resampling via the following steps. First, we randomly select 1 of the 1000 KDE fits. Second, we subset that estimate by the individual's time window, providing the possible KDE values from the full range of time when that person may have died. Third, we randomly select one of the years from that subset and retain the KDE value from that year. Fourth, these steps are repeated 10,000 times per individual. This provides 10,000 distinct KDE values for each individual representing the potential range of the size of the population that existed when they died. Sampling in this fashion helps capture the uncertainty in the demographic reconstruction as the 1000 KDE fits provide the envelope of possibilities and it helps capture the uncertainty in when individuals most likely died by randomly selecting only one year in each iteration. For individuals who are radiocarbon dated, the same process is used as described above except that, when it comes to selecting an individual year to sample the KDE value from, the probability a year is selected is weighted by the probability density from the individual's calibrated radiocarbon date. This weights the selection of values toward higher probability years. Central tendency population size values are then assigned to each individual by taking the average of each individual's 10,000 samples. The result is a proxy of population size per zone during the lifespan of each individual, capturing uncertainty and weighted toward the years in which the individual most likely died.

Prior dates as data research has suggested that taphonomic loss of older radiocarbon dateable material has the potential to impact demographic reconstructions using radiocarbon dates<sup>137</sup>. In the Central Andes, scholars have decided to employ taphonomic correction (or not) based upon research question and sampling area<sup>16,29,40,55</sup> and have suggested that while corrected and uncorrected datasets are generally in agreement, some differences can occur<sup>55</sup>. To address uncertainty as to whether corrected or uncorrected reconstructions are more appropriate, the complete analysis here is run twice, once each for uncorrected and corrected demographic reconstructions. To

generate the corrected reconstructions, we apply the global taphonomic correction equation from Surovell et al.<sup>137</sup> to each of the 1000 KDE fits and then implement the same sampling steps listed above for each individual. This provides both a corrected and uncorrected demographic estimate based on 10,000 iterated samplings for each individual. For our analysis, these relative measures of population enable us to evaluate if variation in population size is an influential factor for predicting diet.

**Statistical analyses.** All analyses are conducted in the R statistical environment<sup>136</sup>, with complete code to replicate the analyses available in Supplementary Material 1.

We visualize  $\delta^{15}\text{N}$  and  $\delta^{13}\text{C}$  trends across each of the elevation categories over time using iterated generalized additive model (GAM) fits. GAMs estimate non-linear fits between response and predictor variables using splines, with smoothing parameters estimated using generalized cross validation<sup>138,139</sup>. We construct GAMs for each elevation zone for both  $\delta^{15}\text{N}$  and  $\delta^{13}\text{C}$ , predicting the isotope value with an individual's date. To address uncertainty in when individual's date to, we construct a series of 10,000 GAMs for each of our six elevation and isotope combinations (60,000 total). First, we take each individual and resample their date window 10,000 times, obtaining 10,000 individual year estimates per person, with each estimate representing a unique run. For radiocarbon dated individuals, we weight the probability of selecting a year from their calibrated date window based upon the probability density from the calibration. We then subset the data to obtain datasets of individuals for coastal, mid-elevation, and highland  $\delta^{15}\text{N}$  and do the same for  $\delta^{13}\text{C}$ , with each individual possessing 10,000 dates. For each combination of elevation zone and isotope, we fit one GAM per unique run (10,000 runs), so each individual has a single year estimate in each of the GAMs. Each unique GAM fit is used to predict the isotope value over the study period. We then aggregate the 10,000 GAMs per elevation zone and isotope, generating a mean predicted value, mean standard error, and mean standard deviation per temporal observation. These data are used to fit a central tendency line with 95% confidence intervals for each combination (Fig. 3).

We also visualize the trends over time using 95% confidence interval ellipses plots broken out by time period, elevation zone, and isotope (Supplementary Fig. 1 and Supplementary Material 2) using the **SIBER** package<sup>140</sup> in R.

We assess the correlation of climate and relative population size with diet in the Central Andes empirically using random forest (RF)<sup>141</sup> regression via the **ranger** package<sup>142</sup> in R. R code for replicating the complete analysis is provided in Supplementary Material 1. RF regression is a machine learning method that employs an ensemble approach to generate mean prediction of a dependent variable when relationships may be non-linear and the number of interactions between variables may be large. RFs subsample predictor variables to prevent over-reliance on any single predictor and control for complex interaction between variables, allowing us to parse out the influence of climatic and demographic variables separately. Here, four regression models are generated, one each for  $\delta^{13}\text{C}\text{‰}$  and  $\delta^{15}\text{N}\text{‰}$  using taphonomically uncorrected or corrected demographic estimates. For each set of individuals with  $\delta^{15}\text{N}\text{‰}$  and those with  $\delta^{13}\text{C}\text{‰}$ , we assess the collinearity of predictor variables by calculating the strength of correlation between pairwise comparisons. Variables with correlation strengths less than 0.70<sup>143</sup> are deemed viable for inclusion in the models, but any greater than 0.70 present potential co-variance problems, although RF analysis can handle such highly correlated variables<sup>141</sup>. For both the  $\delta^{13}\text{C}\text{‰}$  and  $\delta^{15}\text{N}\text{‰}$  sets of individuals, all variables are under the 0.70 correlation threshold and we therefore include all variables in the models.

RF models are evaluated with prediction errors and the standard deviation of residuals from tenfold model cross-validation using the **spm** package<sup>144</sup>. We then check for temporal and spatial autocorrelation in the model residuals using acf plots and evaluating the expected vs observed Moran's I values on an inverse distance matrix calculated using the **ape** package<sup>145</sup> in R.

To estimate the influence of climate versus population size, we calculate the effect size of each covariate. As random forest fits non-linear responses, we use the centered partial dependency function to estimate effect size. In the absence of interactions between variables, the overall predicted response ( $F(X)$ ) is the sum of individual partial dependencies for each covariate<sup>146</sup>. We therefore define effect size for variable  $j$  as:

$$\text{Effect}(x_j) = \sqrt{\sum_{i=0}^n \hat{F}(x_j)}$$

where  $\hat{F}(x_j)$  is the centered partial dependency for  $j$ . A proportional effect size can then be calculated as the ratio of this value to the overall predicted response.

As there are interactions between our variables, we calculate Friedman's  $H$  statistic, or the proportion of each variable's effect resulting from its interaction with the other variables<sup>146</sup> within each elevation category:

$$H = \sqrt{\sum_{i=0}^n F(X) - \hat{F}(x_j) - \hat{F}(x_{-j})} / \sqrt{\sum_{i=0}^n F(X)}$$

Calculation of this statistic is done using the **iml** package in R<sup>147</sup>. Friedman's  $H$  statistic allows us, for each variable, to deduct the proportion of the effect due to the interaction with other variables to get the independent effect size of each predictor variable.

Effect sizes and  $H$  statistic values are estimated through 100 iterated simulations wherein, for each iteration, we randomly sample 100 values of variable  $j$  and assign those values to each individual in the dataset. This results in 100 copies of each individual where the difference is the value of variable  $j$ . We then predict the isotopic value for each individual using the RF model. Using the resultant output, we calculate the total response of the isotope

( $\delta^{13}\text{C}\text{‰}$  or  $\delta^{15}\text{N}\text{‰}$ ), the partial dependence of the isotope to the chosen predictor variable, the combined partial dependence to all other predictor variables, Friedman's *H* statistic, explained sum of squares, explained sum of squares deducted by Friedman's *H* statistic, and the explained sum of squares as a percentage of the total variability in the isotope for the chosen variable. The results of each of the 100 iterations are then averaged together to provide an estimate of the effect size of each variable within each elevation category. We take the square root of the effect sizes to return the values into per mil (‰) space. We then divide the square root of the explained sum of square effect sizes by the total variability to obtain the percent of the variability explained by each of our predictor variables, enabling comparison within models. We compare cumulative climate with demographic effects to evaluate the relative influence of climate versus population size on dietary variation. This comparison removes the influence of the interactions between variables to enable evaluation of the independent effects.

**Data limitations.** Our isotope dataset is the largest compiled in the Americas and we are confident in its usefulness for characterizing differentiation in individuals' diets over time in the Central Andes. However, as with all data, there are several limitations. First, the isotope record is not a random sample and possesses some unevenness in its spatial and temporal distribution (Supplementary Figs. S2 & S3). Numerous individuals represent the past approximately 2000 yBP (Supplementary Figs. S2, S3) on the coasts and mid-elevations, but fewer predate this point in time or are from the highlands. This makes us more confident in the coastal, mid-elevation, and post-2000 yBP patterns than the pre-2000 yBP and highland patterns. Further, we rely on broadly local estimates of climate and population, which might not represent the circumstances experienced by highly mobile individuals e.g.,<sup>148</sup>. Finally,  $\delta^{15}\text{N}$  and  $\delta^{13}\text{C}$  capture broad dietary differences between resources such as marine and camelid protein or maize (C4) and potatoes (C3). Changes in diet that are not shifts between C3 and C4, such as shifting from one C3 resource to another C3 resource, in response to climatic or population change, may not be reflected in this data. Future work combining additional aspects of diet such as isotopic signatures of local food items by site, primary and secondary dietary inputs from zooarchaeological and palaeobotanical analyses, adding additional highland or early individuals, and incorporating additional isotopes such as  $\delta^{18}\text{O}$  and  $\delta^{34}\text{S}$  or  $\delta^{13}\text{C}$  from enamel or bone apatite may reveal unseen patterns and better enable evaluation of what direct food items contribute to changes. The data also are potentially limited by a lack of local isotope baselines for all samples for generating corrections, uncontrolled impacts from possible fertilization, and by implementation of different foddering practices for camelids, even within sites<sup>80,85–87</sup>. Each of these limitations highlights potentially productive avenues for future research.

Additionally, our climate data are spatially coarse-grained simulations ( $2.5 \times 2.5^\circ$  lat/long) which do not capture highly localized variation in climatic patterns. However, the Trace21ka simulation provides complete coverage in time and space across the study area. Attempting to instead rely on proxy records would severely limit the spatial and temporal variation we employ as available proxy records are inconsistent in the temporal periods they cover and would provide less spatial resolution. Given the spatial distribution of the individuals included in this analysis, the Trace21ka simulations provide spatial and temporal variation (Figs. 1, 2) across the geographic and temporal spread of the entire study area. The data we use provide a conservative estimate of the influence of climate and population given the spatial variation they do average over. Future work employing downscaling of GCM variables and/or generating and using more regional paleoclimate proxy data may improve the accuracy of the analyzed relationships.

Finally, using dates as data employs a range of assumptions which may impact the accuracy of demographic trend recreations and KDEs, like summed probability distributions, are flawed proxies for past population estimates<sup>149</sup>. To address some of the flaws, researchers employing this approach have proposed applying taphonomic corrections<sup>137</sup>, binning dates to control for the influence of well-dated sites<sup>125,131</sup>, and separating regions and/or time periods with significantly different human behaviors for analysis<sup>123</sup>. As it is unclear whether taphonomically correcting KDEs is more or less accurate than leaving them uncorrected in this case, we run the analysis twice, once for each the corrected and uncorrected data. To deal with sampling biases, we 'bin' the dates using hierarchical clustering with a 200 year cut-off value and we separate KDEs into elevation zones to attempt to compare KDE fluctuations within contexts likely sharing similar behaviors. We attempt to deal with the assumptions in the best way possible, though violations of them will lower the correlative and interpretive power of our demographic variable. Despite these limitations, KDEs are useful in this case as they provide an estimate of the relative variation in population size and our analysis is concerned with such relative, rather than absolute, differences. While our proxies for both climate/environment and demography are imperfect, they remain useful for assessing the relative influence of these two factors on diet.

## Data availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or Supplementary materials.

Received: 11 August 2021; Accepted: 13 January 2022

Published online: 07 February 2022

## References

1. Kintigh, K. W. *et al.* Grand challenges for archaeology. *Proc. Natl. Acad. Sci. USA* **111**, 879–880 (2014).
2. Bird, D. W. & O'Connell, J. F. Behavioral ecology and archaeology. *J. Archaeol. Res.* **14**, 143–188 (2006).
3. Codding, B. F. & Bird, D. W. Behavioral ecology and the future of archaeological science. *J. Archaeol. Sci.* **56**, 9–20. <https://doi.org/10.1016/j.jas.2015.02.027> (2015).
4. Kennett, D. J. & Winterhalder, B. *Behavioral Ecology and the Transition to Agriculture* (University of California Press, 2006).
5. Barsbai, T., Lukas, D. & Ponderfer, A. Local convergence of behavior across species. *Science* **371**, 292–295 (2021).

6. Coltrain, J. B. & Leavitt, S. W. Climate and diet in Fremont prehistory: economic variability and abandonment of maize agriculture in the Great Salt Lake Basin. *Am. Antiquity* **56**, 453–485 (2002).
7. Seltzer, G. O. & Hastorf, C. A. Climatic change and its effect on prehispanic agriculture in the central Peruvian Andes. *J. Field Archaeol.* **17**, 397–414 (1990).
8. Schwindt, D. M. *et al.* The social consequences of climate change in the central Mesa Verde region. *Am. Antiquity* **81**, 74–96 (2016).
9. Coddling, B. F. *et al.* Socioecological dynamics structuring the spread of farming in the North American Basin-Plateau Region. *Environ. Archaeol.* **56**, 1–13 (2021).
10. Winterhalder, B. & Goland, C. On population, foraging efficiency, and plant domestication. *Curr. Anthropol.* **34**, 710–715 (1993).
11. Boserup, E. *The Conditions of Agricultural Growth: The Economics of Agrarian Change Under Population Pressure* (Aldine, 1965).
12. Weitzel, E. M. & Coddling, B. F. Population growth as a driver of initial domestication in Eastern North America. *R. Soc. Open Sci.* **3**, 160319 (2016).
13. Morgan, C. Is it intensification yet? Current archaeological perspectives on the evolution of hunter-gatherer economies. *J. Archaeol. Res.* **23**, 163–213 (2015).
14. Malthus, T. R. *An Essay on the Principle of Population, as it Affects the Future Improvement of Society, with Remarks on the Speculations of Mr. Godwin, M. Condorcet, and Other Writers* (The Lawbook Exchange, 1798).
15. Venkataraman, V. V., Kraft, T. S., Dominy, N. J. & Endicott, K. M. Hunter-gatherer residential mobility and the marginal value of rainforest patches. *Proc. Natl. Acad. Sci. USA* **114**, 3097–3102 (2017).
16. Roscoe, P., Sandweiss, D. H. & Robinson, E. Population density and size facilitate interactive capacity and the rise of the state. *Philos. Trans. R. Soc. B* **376**, 20190725 (2021).
17. McMichael, A. J. Insights from past millennia into climatic impacts on human health and survival. *Proc. Natl. Acad. Sci. USA* **109**, 4730–4737 (2012).
18. Larsen, C. S. *et al.* Bioarchaeology of Neolithic Çatalhöyük reveals fundamental transitions in health, mobility, and lifestyle in early farmers. *Proc. Natl. Acad. Sci. USA* **116**, 12615–12623 (2019).
19. Kennett, D. J. & Marwan, N. Climatic volatility, agricultural uncertainty, and the formation, consolidation and breakdown of preindustrial agrarian states. *Philos. Trans. R. Soc. A* **373**, 20140458 (2015).
20. 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).
21. McLaughlin, T. R., Gómez-Puche, M., Cascalleira, J., Bicho, N. & Fernández-López de Pablo, J. Late Glacial and Early Holocene human demographic responses to climatic and environmental change in Atlantic Iberia. *Philos. Trans. R. Soc. B* **376**, 20190724 (2021).
22. Williams, A. N. *et al.* A continental narrative: Human settlement patterns and Australian climate change over the last 35,000 years. *Quatern. Sci. Rev.* **123**, 91–112 (2015).
23. Bevan, A. *et al.* Holocene fluctuations in human population demonstrate repeated links to food production and climate. *Proc. Natl. Acad. Sci. USA* **114**, E10524–E10531 (2017).
24. Contreras, D. A. Landscape and environment: insights from the prehispanic Central Andes. *J. Archaeol. Res.* **18**, 241–288 (2010).
25. Sandweiss, D. H. & Richardson, J. B. in *The handbook of South American archaeology* Ch. 6, 93–104 (Springer, 2008).
26. Kirch, P. V. The evolution of sociopolitical complexity in prehistoric Hawaii: an assessment of the archaeological evidence. *J. World Prehist.* **4**, 311–345 (1990).
27. Pulgar Vidal, J. *Geografía del Perú: Las ocho regiones naturales del Perú*. 8th edn, (Editorial Universo, 1981).
28. Gerdau-Radonić, K. *et al.* Diet in Peru's pre-Hispanic central coast. *J. Archaeol. Sci. Rep.* **4**, 371–386 (2015).
29. Riris, P. & Arroyo-Kalin, M. Widespread population decline in South America correlates with mid-Holocene climate change. *Sci. Rep.* **9**, 1–10 (2019).
30. Carneiro, R. L. A Theory of the Origin of the State. *Science* **169**, 733–738 (1970).
31. Silverman, H. & Isbell, W. H. *Handbook of South American Archaeology* (Springer, 2008).
32. Aufderheide, A. C., Muñoz, I. & Arriaza, B. Seven Chinchorro mummies and the prehistory of northern Chile. *Am. J. Phys. Anthropol.* **91**, 189–201 (1993).
33. Moseley, M. E. *The Maritime Foundations of Andean Civilization* (Cummings Publishing Company, 1975).
34. Haas, R. *et al.* Humans permanently occupied the Andean highlands by at least 7 ka. *R. Soc. Open Sci.* **4**, 170331 (2017).
35. Parsons, J. R., Hastings, C. M. & Matos, R. Rebuilding the state in highland Peru: Herder-cultivator interaction during the Late Intermediate period in the Tarama-Chinchaycocha region. *Latin American Antiquity*, 317–341 (1997).
36. Troll, C. in *Colloquium Geographicum (Univ. Bonn)*. 15–56.
37. Browman, D. L. in *Arid land use strategies and risk management in the Andes: A regional anthropological perspective* (ed David L Browman) 121–149 (Westview Press, 1987).
38. Pearsall, D. M. in *The handbook of South American archaeology* (eds Helaine Silverman & William H Isbell) Ch. 7, 105–120 (Springer Science & Business Media, 2008).
39. Murra, J. V. in *Visita de la Provincia de Leon de Huanuco en 1562 por Inigo Ortiz de Zuniga* (ed John V Murra) 427–476 (Universidad Nacional Hermilio Valdizan, 1972).
40. Riris, P. Dates as data revisited: A statistical examination of the Peruvian preceramic radiocarbon record. *J. Archaeol. Sci.* **97**, 67–76 (2018).
41. Dillehay, T. D. & Kolata, A. L. Long-term human response to uncertain environmental conditions in the Andes. *Proc. Natl. Acad. Sci. USA* **101**, 4325–4330 (2004).
42. Haas, J. & Creamer, W. Crucible of Andean civilization: The Peruvian coast from 3000 to 1800 BC. *Curr. Anthropol.* **47**, 745–775 (2006).
43. Hastorf, C. A. & Johannessen, S. Pre-Hispanic political change and the role of maize in the Central Andes of Peru. *Am. Anthropol.* **95**, 115–138 (1993).
44. Weng, C. *et al.* Deglaciation and Holocene climate change in the western Peruvian Andes. *Quaternary Res.* **66**, 87–96 (2006).
45. Quinn, W. H. The Large-Scale ENSO Event, The El Niño and Other Important Regional Features. *Bull. Inst. fr. études an dines* **22**, 13–34 (1993).
46. Sandweiss, D. H. *et al.* Archaeological climate proxies and the complexities of reconstructing Holocene El Niño in coastal Peru. *Proc. Natl. Acad. Sci. USA* **117**, 8271–8279 (2020).
47. Goodbred, S. L. Jr., Dillehay, T. D., Mora, C. G. & Sawakuchi, A. O. Transformation of maritime desert to an agricultural center: Holocene environmental change and landscape engineering in Chicama River valley, northern Peru coast. *Quaternary Sci. Rev.* **227**, 106046 (2020).
48. Thompson, L. G., Mosley-Thompson, E., Bolzan, J. F. & Koci, B. R. A 1500-year record of tropical precipitation in ice cores from the Quelccaya ice cap, Peru. *Science* **229**, 971–973 (1985).
49. Binford, M. W. *et al.* Climate variation and the rise and fall of an Andean civilization. *Quaternary Res.* **47**, 235–248 (1997).
50. Thompson, L. G., Mosley-Thompson, E. & Henderson, K. A. Ice-core palaeoclimate records in tropical South America since the Last Glacial Maximum. *J. Quat. Sci. Publ. Quat. Res. Assoc.* **15**, 377–394 (2000).
51. Sandweiss, D. H., Solís, R. S., Moseley, M. E., Keefer, D. K. & Orloff, 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).

52. Shimada, I., Schaaf, C. B., Thompson, L. G. & Mosley-Thompson, E. Cultural impacts of severe droughts in the prehistoric Andes: Application of a 1,500-year ice core precipitation record. *World Archaeol.* **22**, 247–270 (1991).
53. Grosjean, M., Núñez, L., Cartajena, I. & Messerli, B. Mid-Holocene climate and culture change in the Atacama Desert, northern Chile. *Quat. Res.* **48**, 239–246 (1997).
54. Fehren-Schmitz, L. *et al.* Climate change underlies global demographic, genetic, and cultural transitions in pre-Columbian southern Peru. *Proc. Natl. Acad. Sci. USA* **111**, 9443–9448 (2014).
55. 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).
56. Batai, K. & Williams, S. R. Mitochondrial variation among the Aymara and the signatures of population expansion in the central Andes. *Am. J. Hum. Biol.* **26**, 321–330 (2014).
57. Perez, S. I., Postillone, M. B. & Rindel, D. Domestication and human demographic history in South America. *Am. J. Phys. Anthropol.* **163**, 44–52 (2017).
58. Moseley, M. E. *The Incas and their ancestors: the archaeology of Peru*. Revised edn, (Thames and Hudson, 2001).
59. Rivera, M. A. The prehistory of northern Chile: A synthesis. *J. World Prehist.* **5**, 1–47 (1991).
60. Tung, T. A., Dillehay, T. D., Feranec, R. S. & DeSantis, L. R. Early specialized maritime and maize economies on the north coast of Peru. *Proc. Natl. Acad. Sci. USA* **117**, 32308–32319 (2020).
61. Solis, R. S., Haas, J. & Creamer, W. Dating Caral, a preceramic site in the Supe Valley on the central coast of Peru. *Science* **292**, 723–726 (2001).
62. Pozorski, T. & Pozorski, S. Early complex society on the north and central Peruvian coast: new archaeological discoveries and new insights. *J. Archaeol. Res.* **26**, 353–386 (2018).
63. Aldenderfer, M. Preludes to power in the highland Late Preceramic Period. *Archeol. Pap. Am. Anthropol. Assoc.* **14**, 13–35 (2004).
64. Cuéllar, A. M. The Archaeology of Food and Social Inequality in the Andes. *J. Archaeol. Res.* 123–174 (2013).
65. Kellner, C. M. & Schoeninger, M. J. Wari's imperial influence on local Nasca diet: the stable isotope evidence. *J. Anthropol. Archaeol.* **27**, 226–243 (2008).
66. Stanish, C., Tantaleán, H. & Knudson, K. Feasting and the evolution of cooperative social organizations circa 2300 BP in Paracas culture, southern Peru. *Proc. Natl. Acad. Sci. USA* **115**, E6716–E6721 (2018).
67. Bongers, J. L. *et al.* Integration of ancient DNA with transdisciplinary dataset finds strong support for Inca resettlement in the south Peruvian coast. *Proc. Natl. Acad. Sci. USA* **117**, 18359–18368 (2020).
68. Isbell, W. H. in *The handbook of South American archaeology* (eds Helaine Silverman & William H Isbell) Ch. 37, 731–759 (Springer Science & Business Media, 2008).
69. D'Altroy, T. N. & Hastorf, C. A. *Empire and Domestic Economy* (Kluwer Academic Publishers, 2001).
70. DeFrance, S. D. Fishing specialization and the inland trade of the Chilean jack mackerel or jurel, *Trachurus murphyi*, in far southern Peru. *Archaeol. Anthropol. Sci.* **13**, 1–13 (2021).
71. Staller, J. E. in *Histories of maize: Multidisciplinary approaches to the prehistory, linguistics, biogeography, domestication, and evolution of maize* (eds John Staller, Robert H Tykot, & Bruce F Benz) 449–467 (Routledge, 2006).
72. Sandweiss, D. H. in *The handbook of south American archaeology* (eds Helaine Silverman & William H Isbell) 145–156 (Springer, 2008).
73. Prieto, G. & Sandweiss, D. H. *Maritime Communities of the Ancient Andes: Gender, Nation, and Popular Culture*. (University Press of Florida, 2019).
74. King, C. L. *et al.* Marine resource reliance in the human populations of the Atacama Desert, northern Chile—A view from prehistory. *Quatern. Sci. Rev.* **182**, 163–174 (2018).
75. Tieszen, L. L., Iversen, E. & Matzner, S. in *Proceedings of the First World Congress on Mummy Studies*. 427–441 (Museo Arqueológico y Etnográfico de Tenerife Santa Cruz de Tenerife).
76. Pestle, W. J., Torres-Rouff, C., Gallardo, F., Ballester, B. & Clarot, A. Mobility and exchange among marine hunter-gatherer and agropastoralist communities in the Formative Period Atacama Desert. *Curr. Anthropol.* **56**, 121–133 (2015).
77. Tieszen, L. L. & Chapman, M. in *Actas del I Congreso Internacional de Estudios sobre Momias= Proceedings of the I World Congress on Mummy Studies*. 409–441.
78. Shady, R. in *Histories of maize: multidisciplinary approaches to the prehistory, linguistics, biogeography, domestication and evolution of maize*. Amsterdam: Elsevier (eds John Staller, Robert Tykot, & Bruce Benz) 381–402 (Routledge, 2006).
79. Burger, R. L. & van der Merwe, N. J. Maize and the origin of highland Chavín civilization: an isotopic perspective. *Am. Anthropol.* **92**, 85–95 (1990).
80. Finucane, B., Agurto, P. M. & Isbell, W. H. Human and animal diet at Conchopata, Peru: stable isotope evidence for maize agriculture and animal management practices during the Middle Horizon. *J. Archaeol. Sci.* **33**, 1766–1776 (2006).
81. Zimmerer, K. S. *Changing fortunes: Biodiversity and peasant livelihood in the Peruvian Andes*. (University of California Press, 1996).
82. Finucane, B. C. Maize and sociopolitical complexity in the Ayacucho Valley, Peru. *Curr. Anthropol.* **50**, 535–545 (2009).
83. Hyland, C., Millaire, J. F. & Szpak, P. Migration and maize in the Virú Valley: Understanding life histories through multi-tissue carbon, nitrogen, sulfur, and strontium isotope analyses. *Am. J. Phys. Anthropol.* (2021).
84. Takigami, M. *et al.* Isotopic study of maize exploitation during the Formative Period at Pacopampa, Peru. *Anthropol. Sci.*, 210531, doi:<https://doi.org/10.1537/ase.210531> (2021).
85. Szpak, P., White, C. D., Longstaffe, F. J., Millaire, J.-F. & Sánchez, V. F. V. Carbon and nitrogen isotopic survey of northern Peruvian plants: baselines for paleodietary and paleoecological studies. *PLoS ONE* **8**, e53763 (2013).
86. Szpak, P., Longstaffe, F. J., Millaire, J.-F. & White, C. D. Large variation in nitrogen isotopic composition of a fertilized legume. *J. Archaeol. Sci.* **45**, 72–79 (2014).
87. Szpak, P., Millaire, J.-F., White, C. D. & Longstaffe, F. J. Influence of seabird guano and camelid dung fertilization on the nitrogen isotopic composition of field-grown maize (*Zea mays*). *J. Archaeol. Sci.* **39**, 3721–3740 (2012).
88. Marsteller, S. J., Zolotova, N. & Knudson, K. J. Investigating economic specialization on the central Peruvian coast: A reconstruction of Late Intermediate Period Ychsma diet using stable isotopes. *Am. J. Phys. Anthropol.* **162**, 300–317 (2017).
89. McCool, W. C., Tung, T. A., Coltrain, J. B., Accinelli Obando, A. J. & Kennett, D. J. The character of conflict: A bioarchaeological study of violence in the Nasca highlands of Peru during the Late Intermediate Period (950–1450 CE). *Am. J. Phys. Anthropol.* (2020).
90. Tung, T. A., Miller, M., DeSantis, L., Sharp, E. A. & Kelly, J. in *The archaeology of food and warfare* 193–228 (Springer, 2016).
91. La Lone, M. B. & La Lone, D. E. The Inka State in the Southern Highlands: state administrative and production enclaves. *Ethnohistory*, 47–62 (1987).
92. Finucane, B. C. Mummies, maize, and manure: multi-tissue stable isotope analysis of late prehistoric human remains from the Ayacucho Valley, Peru. *J. Archaeol. Sci.* **34**, 2115–2124 (2007).
93. Janusek, J. W. & Kolata, A. L. Top-down or bottom-up: rural settlement and raised field agriculture in the Lake Titicaca Basin, Bolivia. *J. Anthropol. Archaeol.* **23**, 404–430 (2004).
94. Burger, R. L., Lee-Thorp, J. A. & Van der Merwe, N. in *1912 Yale Peruvian scientific expedition collections from Macchu Picchu: human and animal remains Yale University Publications in Anthropology* 119–137 (Peabody Museum of Natural History, 2003).

95. Vinton, S. D., Perry, L., Reinhard, K. J., Santoro, C. M. & Teixeira-Santos, I. Impact of empire expansion on household diet: the Inka in Northern Chile's Atacama Desert. *PLoS ONE* **4**, e8069 (2009).
96. D'Altroy, T. N. *et al.* Staple finance, wealth finance, and storage in the Inka political economy. *Curr. Anthropol.* **26**, 187–206 (1985).
97. Borda, V. *et al.* The genetic structure and adaptation of Andean highlanders and Amazonians are influenced by the interplay between geography and culture. *Proc. Natl. Acad. Sci. USA* **117**, 32557–32565 (2020).
98. Stanish, C. *et al.* Tiwanaku trade patterns in southern Peru. *J. Anthropol. Archaeol.* **29**, 524–532 (2010).
99. Covey, R. A., Bauer, B. S., Bélisle, V. & Tsesmeli, L. Regional perspectives on Wari state influence in Cusco, Peru (c. AD 600–1000). *J. Anthropol. Archaeol.* **32**, 538–552 (2013).
100. Schoeninger, M. J. & Moore, K. Bone Stable Isotope Studies in Archaeology. *J. World Prehist.* **6**, 247–296 (1992).
101. Knudson, K. J., Peters, A. H. & Cagigao, E. T. Paleodiet in the Paracas Necropolis of Wari Kayan: carbon and nitrogen isotope analysis of keratin samples from the south coast of Peru. *J. Archaeol. Sci.* **55**, 231–243 (2015).
102. Turner, B. L. *et al.* Diet and foodways across five millennia in the Cusco region of Peru. *J. Archaeol. Sci.* **98**, 137–148 (2018).
103. Cadwallader, L. *Investigating 1500 Years of Dietary Change in the Lower Ica Valley, Peru Using an Isotopic Approach* Doctor of Philosophy thesis, University of Cambridge (2013).
104. Alfonso-Durruty, M. P. *et al.* Dietary diversity in the Atacama desert during the Late intermediate period of northern Chile. *Quatern. Sci. Rev.* **214**, 54–67 (2019).
105. Hedges, R. E., Clement, J. G., Thomas, C. D. L. & O'Connell, T. C. Collagen turnover in the adult femoral mid-shaft: Modeled from anthropogenic radiocarbon tracer measurements. *Am. J. Phys. Anthropol.* **133**, 808–816 (2007).
106. Seifert, M. F. & Watkins, B. A. Role of dietary lipid and antioxidants in bone metabolism. *Nutr. Res.* **17**, 1209–1228 (1997).
107. Ambrose, S. H. Preparation and characterization of bone and tooth collagen for isotopic analysis. *J. Archaeol. Sci.* **17**, 431–451 (1990).
108. Schoeninger, M. J., Moore, K. M., Murray, M. L. & Kingston, J. D. Detection of bone preservation in archaeological and fossil samples. *Appl. Geochem.* **4**, 281–292 (1989).
109. Cerling, T. E., Harris, J. M., Ambrose, S. H., Leakey, M. G. & Solounias, N. Dietary and environmental reconstruction with stable isotope analyses of herbivore tooth enamel from the Miocene locality of Fort Ternan, Kenya. *J. Hum. Evol.* **33**, 635–650 (1997).
110. DeNiro, M. J. & Epstein, S. Influence of diet on the distribution of nitrogen isotopes in animals. *Geochim. Cosmochim. Acta* **45**, 341–351 (1981).
111. Lee-Thorp, J. A. & van der Merwe, N. J. Aspects of the chemistry of modern and fossil biological apatites. *J. Archaeol. Sci.* **18**, 343–354 (1991).
112. Bender, M. M. Variations in the <sup>13</sup>C/<sup>12</sup>C ratios of plants in relation to the pathway of photosynthetic carbon dioxide fixation. *Phytochemistry* **10**, 1239–1244 (1971).
113. Pate, F. D. Bone chemistry and paleodiet. *J. Archaeol. Method Theory* **1**, 161–209 (1994).
114. Ambrose, S. H. & DeNiro, M. J. Reconstruction of African human diet using bone collagen carbon and nitrogen isotope ratios. *Nature* **319**, 321–324 (1986).
115. Schoeller, D. A. Isotope fractionation: why aren't we what we eat?. *J. Archaeol. Sci.* **26**, 667–673 (1999).
116. Schulting, R. J. Slighting the sea: stable isotope evidence for the transition to farming in northwestern Europe. *Documenta Praehistorica* **25**, 203–218 (1998).
117. Turner, B. L., Kingston, J. D. & Armelagos, G. J. Variation in dietary histories among the immigrants of Machu Picchu: Carbon and nitrogen isotope evidence. *Chungara, Revista de Antropología Chilena* **42**, 515–534 (2010).
118. Szpak, P., Millaire, J.-F., White, C. D. & Longstaffe, F. J. Small scale camelid husbandry on the north coast of Peru (Virú Valley): insight from stable isotope analysis. *J. Anthropol. Archaeol.* **36**, 110–129 (2014).
119. Liu, Z. *et al.* Evolution and forcing mechanisms of El Niño over the past 21,000 years. *Nature* **515**, 550–553 (2014).
120. Otto-Bliesner, B. L. *et al.* Coherent changes of southeastern equatorial and northern African rainfall during the last deglaciation. *Science* **346**, 1223–1227 (2014).
121. Liu, Z. *et al.* Transient simulation of last deglaciation with a new mechanism for Bølling-Allerød warming. *Science* **325**, 310–314 (2009).
122. Fordham, D. A. *et al.* PaleoView: a tool for generating continuous climate projections spanning the last 21 000 years at regional and global scales. *Ecography* **40**, 1348–1358 (2017).
123. Rick, J. W. Dates as data: an examination of the Peruvian preceramic radiocarbon record. *Am. Antiquity*, 55–73 (1987).
124. Shennan, S. *et al.* Regional population collapse followed initial agriculture booms in mid-Holocene Europe. *Nat. Commun.* **4**, 1–8 (2013).
125. Timpson, A. *et al.* Reconstructing regional population fluctuations in the European Neolithic using radiocarbon dates: a new case-study using an improved method. *J. Archaeol. Sci.* **52**, 549–557 (2014).
126. Ziolkowski, M. S. *Andes: Radiocarbon Database for Bolivia, Ecuador and Peru* (Warsaw University, 1994).
127. Rademaker, K., Bromley, G. R. & Sandweiss, D. H. Peru archaeological radiocarbon database, 13,000–7000 14C BP. *Quat. Int.* **301**, 34–45 (2013).
128. Goldberg, A., Mychajliw, A. M. & Hadly, E. A. Post-invasion demography of prehistoric humans in South America. *Nature* **532**, 232–235 (2016).
129. Robinson, E., Zahid, H. J., Coddling, B. F., Haas, R. & Kelly, R. L. Spatiotemporal dynamics of prehistoric human population growth: radiocarbon 'dates as data' and population ecology models. *J. Archaeol. Sci.* **101**, 63–71 (2019).
130. Brown, W. A. The past and future of growth rate estimation in demographic temporal frequency analysis: biodemographic interpretability and the ascendance of dynamic growth models. *J. Archaeol. Sci.* **80**, 96–108 (2017).
131. Crema, E. R. & Bevan, A. Inference from large sets of radiocarbon dates: software and methods. *Radiocarbon* **63**, 23–39 (2021).
132. Crema, E. R., Bevan, A. & Shennan, S. Spatio-temporal approaches to archaeological radiocarbon dates. *J. Archaeol. Sci.* **87**, 1–9 (2017).
133. Bronk Ramsey, C. Methods for summarizing radiocarbon datasets. *Radiocarbon* **59**, 1809–1833 (2017).
134. Reimer, P. J. & Reimer, R. W. A marine reservoir correction database and on-line interface. *Radiocarbon* **43**, 461–463 (2001).
135. Marsh, E. J. *et al.* IntCal, SHCal, or a mixed curve? Choosing a 14C calibration curve for archaeological and paleoenvironmental records from tropical South America. *Radiocarbon* **60**, 925–940 (2018).
136. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/> (2021).
137. 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).
138. Wood, S. N. *Generalized Additive Models: An Introduction with R*. (Chapman and Hall/CRC, 2017).
139. Wood, S. N. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **73**, 3–36 (2011).
140. Jackson, A. L., Inger, R., Parnell, A. C. & Bearhop, S. Comparing isotopic niche widths among and within communities: SIBER—Stable Isotope Bayesian Ellipses in R. *J. Anim. Ecol.* **80**, 595–602 (2011).
141. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).

142. Wright, M. N. & Ziegler, A. ranger: A fast implementation of random forests for high dimensional data in C++ and R. *J. Stat. Softw.* **77**, 1–17 (2017).
143. Dormann, C. F. *et al.* Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 27–46 (2013).
144. spm: Spatial Predictive Modeling. R package version 1.2.0. <https://CRAN.R-project.org/package=spm> (2019).
145. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528. <https://doi.org/10.1093/bioinformatics/bty633> (2019).
146. Friedman, J. H. & Popescu, B. E. Predictive learning via rule ensembles. *Ann. Appl. Stat.* **2**, 916–954 (2008).
147. Molnar, C., Casalicchio, G. & Bischl, B. iml: An R package for interpretable machine learning. *J. Open Source Softw.* **3**, 786 (2018).
148. Knudson, K. J. & Tung, T. A. Investigating regional mobility in the southern hinterland of the Wari Empire: biogeochemistry at the site of Beringa, Peru. *Am. J. Phys. Anthropol.* **145**, 299–310 (2011).
149. 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).
150. Becker, R. A., Wilks, A. R. & Brownrigg, R. mapdata: Extra map databases. R package version 2.3.0. <https://CRAN.R-project.org/package=mapdata> (2014).
151. Hijmans, R. J. Geographic Data Analysis and Modeling R package version 3.5-2. <https://CRAN.R-project.org/package=raster> (2021).
152. Wessel, P. & Smith, W. H. A global, self-consistent, hierarchical, high-resolution shoreline database version 2.3.7. *J. Geophys. Res. Solid Earth* **101**, 8741–8743. <https://doi.org/10.1029/96JB00104> (1996).
153. Xie, Y. animation: An R package for creating animations and demonstrating statistical methods. *J. Stat. Softw.* **53**, 1–27 (2013).

## Acknowledgements

Kasey Cole (University of Utah), Kate Magargal (University of Utah), Peter Yaworsky (Aarhus University; University of Utah), Kenneth Blake Vernon (University of Utah), and other members of the University of Utah Archaeological Center (UUAC) offered valuable insights and comments on various phases of the analysis and manuscript. Stephanie Gruver (University of Florida) contributed site coordinates for some of the radiocarbon dates used in the generation of the population proxy. Funding for this research was provided by the Center for Latin American Studies, the Global Change and Sustainability Center, and the Vice President for Research Office at the University of Utah.

## Author contributions

K.M.W., B.F.C., and D.A.C. designed the research. K.M.W., W.C.M., A.M.H., N.Z.W., J.B.C., P.J.S., R.F.L., and D.A.C. collected the data. K.M.W., B.F.C., S.C.B., and D.A.C. analyzed the data. K.M.W. wrote the main text and all authors edited the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-05774-y>.

**Correspondence** and requests for materials should be addressed to K.M.W.

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

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022