Productivity, biodiversity, and pathogens influence the global hunter-gatherer population density

Miikka Tallavaara,1,* Jussi T. Eronen,2,3 and Miska Luoto

*Department of Geosciences and Geography, University of Helsinki, FI-00014 Helsinki, Finland; and 2BIOS Research Unit, FI-00170 Helsinki, Finland

The environmental drivers of species distributions and abundances are at the core of ecological research. However, the effects of these drivers on human abundance are not well-known. Here, we report how net primary productivity, biodiversity, and pathogen stress affect human population density using global ethnographic hunter-gatherer data. Our results show that productivity has significant effects on population density globally. The most important direct drivers, however, depend on environmental conditions: biodiversity influences population density exclusively in low-productivity regions, whereas pathogen stress does so in high-productivity regions. Our results also indicate that subtropical and temperate forest biomes provide the highest carrying capacity for hunter-gatherer populations. These findings document that environmental factors play a key role in shaping global population density patterns of preagricultural humans.

Questions related to the environmental drivers of species distributions and abundances have been at the core of ecology throughout its history. However, these questions have rarely been addressed for preindustrial human populations, despite the fact that these populations provide a rare opportunity to investigate drivers of the distribution and abundance of a single species along global environmental gradients. Recent archaeological studies have correlated long-term human population dynamics and major dispersal and range shift events with climate changes (1–5), but their results are mixed regarding the significance of climate in shaping human population processes. Ethnographic data provide finer-scale information about preindustrial human populations and their environments, potentially allowing more conclusive analyses. However, even for hunter-gatherers, ethnographic studies about the ecological determinants of population patterns are relatively few (6–9), despite hunter-gatherers’ close dependence on their immediate environment and wild resources.

Here, we use global ethnographic hunter-gatherer data (10, 11) to explore the effects of key environmental variables on human population densities. These variables are net primary productivity (NPP) calculated as a function of the annual mean temperature and annual precipitation; biodiversity as surrogated by a combination of mammal, bird, and vascular plant richness; and environmental pathogen stress as a combination of pathogen richness and the severity of exposure (Materials and Methods and Fig. S1). In previous work, only the effects of productivity on human population densities have been assessed. These studies suggest that both primary and secondary productivity have, at least regionally, positive effects on hunter-gatherer population density as well as on population home ranges (6–9). Such positive effects are expected, because hunter-gatherers access food directly from their surroundings, which vary widely in energy availability (12). Although hunter-gatherers depend on the productivity of wild plant and animal species, they appropriate only a small fraction of the production and in this regard, more resemble other land mammals of their size than they do humans of modern industrialized societies (13).

However, the amount of available food may not be all that matters in controlling hunter-gatherer abundance, as biodiversity can also play a role. Biodiversity influences ecosystem stability, such that with increasing levels of diversity, the temporal variance of ecosystem productivity tends to decrease and the temporal mean of ecosystem productivity tends to increase (14, 15). A current consensus is that higher biodiversity can enhance the temporal stability of aggregate ecosystem properties, such as biomass and productivity, but its effect on stability of individual populations remains controversial (15, 16). The stabilizing effect of biodiversity may come through processes like asynchrony in the responses of species to environmental variation (17). For hunter-gatherers, increased stability of ecosystem-level biomass production caused by higher biodiversity decreases subsistence-related risk, which can positively affect hunter-gatherer population densities.

In contrast to the effects of productivity and biodiversity, the effects of pathogens on hunter-gatherer abundance are negative. A well-known example is the demographic shock among Native Americans caused by infectious diseases after European contact (18, 19). Beyond the effects of such dramatic epidemiological events, we postulate that the background level of pathogen stress specific to each environment also influences human population density. Studies in animal ecology have shown that pathogens have significant effects on population density (20, 21). Here, we show that human population density is similarly affected by pathogen stress.

Significance

Because of complex cumulative culture, human populations are often considered to be divorced from the environment and not be under the same ecological forcing as other species. However, this study shows that key environmental parameters net primary productivity, biodiversity, and environmental pathogen stress have strong influence on the global pattern of hunter-gatherer population density. Productivity and biodiversity exert the strongest influence in high and midlatitudes, whereas pathogens become more important in tropics. The most suitable conditions for preagricultural humans are found in temperate and subtropical biomes. Our results show that cultural evolution has not freed human hunter-gatherers from strong biotic and abiotic forcing.

Author contributions: M.T. and M.L. designed research; M.T. performed research; M.T. and M.L. contributed new reagents/analytic tools; M.T., J.T.E., and M.L. analyzed data; and M.T., J.T.E., and M.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. R.L.K. is a guest editor invited by the Editorial Board.

Published under the PNAS license.

Data deposition: The data and scripts reported in this paper have been deposited in the Zenodo repository, https://doi.org/10.5281/zenodo.1069786.

To whom correspondence should be addressed. Email: miikka.tallavaara@helsinki.fi.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1715638115/-/DCSupplemental.
can have negative effects on hosts’ abundance (20–22), and we expect that this applies to human populations as well.

Previously, Binford (10) has suggested that resource abundance and stability as well as pathogens would have played an important role in constraining hunter-gatherer population growth. However, Binford (10) does not analyze this hypothesis with the data, whereas our study presents a test of the joint effect of NPP, biodiversity, and pathogens on hunter-gatherer populations.

We use structural equation modeling to reveal the potentially hierarchical effects of NPP, biodiversity, and environmental pathogen stress (23) (Materials and Methods). This method allows us to explore the network of variables and thus, the indirect effects of the explanatory variables. In addition to its direct effect on hunter-gatherer population density, NPP may have an indirect effect mediated by biodiversity. Although the actual mechanism is still debated, plant and animal richness is related to productivity, and these relationships are predominantly positive and monotonic (24, 25). Furthermore, pathogen richness follows the same latitudinal pattern as biodiversity in general, and this pattern is most likely related to temperature and precipitation (26). Thus, climate-based NPP may also have an indirect effect on population density through pathogen stress. However, human pathogen richness is also related to the richness of alternative hosts (27). Consequently, the effects of climate-based NPP on population density can be simultaneously mediated by both biodiversity and pathogen stress. This hypothesized network of variables is illustrated in Fig. 1.

In addition, we provide a geographical interpretation of environmental effects on population density by modeling a distribution map of the relative contributions of the limiting effects of NPP, biodiversity, and pathogen stress on global hunter-gatherer population density. We also model a potential distribution of global hunter-gatherer population density itself as a function of the three environmental variables. These maps are based on the combination of two regression modeling techniques and are projections to current environmental conditions assuming that the globe is populated by hunter-gatherers only (Materials and Methods).

**Results**

Visual inspections and breakpoint analysis of the bivariate relationship between climate-based NPP and hunter-gatherer population density indicate an NPP threshold of ~1,360 g/m² per year, above which the population density becomes largely independent of additional increases in annual productivity (Fig. 2A). However, above the same threshold, NPP exhibits a strong predictive power of pathogen stress (Fig. 2B). This raises the question as to whether the leveling off of hunter-gatherer population density in high-productivity environments is related to increased pathogen stress.

To further analyze these potentially complex relationships between variables, we used structural equation modeling. To ease interpretability of the direct and indirect effects, we built separate models with identical structure (Fig. 1) for the low- and high-productivity environments delineated by the NPP threshold identified in the breakpoint analysis.

The structural equation modeling results strongly indicate different constraints between the low- and high-productivity environments. In the low-productivity environments, biodiversity has the strongest direct effect on hunter-gatherer population density followed by NPP, whereas the effect of pathogen stress is not statistically significant (Fig. 3A, Fig. S2A, and Table S1). In the high-productivity regions, however, the importance of the explanatory variables is reversed. Here, pathogen stress has the strongest effect, and NPP still has relatively high direct and indirect effects, but biodiversity is no longer significant (Fig. 3B, Fig. S2B, and Table S1).

Because of its indirect effect through biodiversity, NPP has the strongest total effect (combined direct and indirect effects) on population density in low-productivity regions (Table S1). In high-productivity regions, the total effect of NPP is weaker than the direct effect of pathogens (Table S1).

As a cross-check to the structural equation modeling results regarding the direct controls of hunter-gatherer population density, we fit a linear model to the whole data (i.e., the data were not split into low- and high-productivity environments as in

**Fig. 1.** Structural equation metamodel showing hypothesized expectations about the relationships among NPP, biodiversity, environmental pathogen stress, and hunter-gatherer population density. Blue arrows indicate positive effects, and red arrows indicate negative effects.

**Fig. 2.** Contrasting effects of NPP on hunter-gatherer population density and pathogen stress. (A) Relationship between NPP and hunter-gatherer population density. Bubble size reflects the magnitude of pathogen stress. A bivariate piecewise regression model (black curve) explains 44% of the variation in population density. (B) Relationship between NPP and pathogen stress. Bubble size reflects hunter-gatherer population density. A bivariate piecewise regression model (black curve) explains 29% of the variation in pathogen stress. A and B also show linear model breakpoints, their SEs, and the mean of these breakpoints (1,360 g/m² per year), which is used to define the low- (n = 234) and high-productivity (n = 66) environments. In addition, A and B show coefficients of NPP and their P values in regressions between NPP, population density, and pathogen stress in low- and high-productivity environments.
the structural equation modeling). In this model, we included interaction terms between biodiversity and NPP and between pathogen stress and NPP. The model shows the same pattern of direct effects of biodiversity and pathogen stress on population density as the structural equation modeling (Fig. 3 C and D, Fig. S3, and Table S2): the effect of biodiversity decreases as NPP increases, and the effect of pathogens becomes stronger as NPP increases. This strongly supports our structural equation modeling results regarding the different controls of population density in low- and high-productivity environments.

The modeled geographical distribution of the relative limiting effects of NPP, biodiversity, and pathogens on hunter-gatherer population density suggests that, among the three variables, environmental pathogen stress has the strongest limiting effect on population density in the tropics and subtropics (Fig. 4 A and E). In contrast, NPP and biodiversity constrain hunter-gatherer population density in mid- and high latitudes (Fig. 4 A, C, and D). These two factors tend to be colimiting, especially in high latitudes. NPP assumes a more dominant role in North America and in northeastern Eurasia, whereas biodiversity assumes a more dominant role in dry regions and in midlatitude Eurasia. (Fig. 4 A, C, and D).

Given the effects of NPP, biodiversity, and pathogen stress, the potential global distribution of hunter-gatherers highlights previously known areas of high hunter-gatherer abundance, such as the western coast of North America and the eastern and northeastern coasts of Australia (10, 11) (Fig. 4B and Fig. S4). However, the map also suggests a number of areas with potentially high hunter-gatherer abundances that have been occupied by agricultural populations from at least the late mid-Holocene onward and that are not well-represented in the historical ethnographic data on hunter-gatherers (Fig. 4B and Fig. S4). These areas include southwestern Mexico, southeastern North America, Western Europe, Southeast Asia, and Papua New Guinea. The highest predicted mean population densities are in temperate oceanic forest biome followed by subtropical humid forests, temperate continental forests, and tropical rainforests (Fig. 4F). However, the differences between these highly suitable biomes are small. When interpreting these results, one has to keep in mind that the global projection takes into account only the effects of three variables that reflect the influence of terrestrial environments on hunter-gatherers. Many coastal hunter-gatherer populations have been dependent on aquatic resources, especially in high latitudes, where terrestrial productivity is relatively low (10, 11). Thus, it is possible that the global projection of population density would change if productivity of aquatic environment is taken into account.

Discussion and Conclusion

Our results show that NPP has significant positive effects on hunter-gatherer abundance globally. However, there is no single determinant of the large-scale variation in abundance. Rather, the most influential direct drivers of population density are context-dependent. Biodiversity affects population density mostly in the low-productivity environments, whereas pathogen stress, which has insignificant effects in areas of low productivity, is a crucial constraint on population density in the high-productivity settings, especially the tropics. The insignificant effect of biodiversity on hunter-gatherer population density in high-productivity environments is not totally unexpected, as ecosystem properties tend to have a saturating response to increasing species richness (14, 15). In high-productivity settings, which often coincide with high biodiversity, increasing diversity does not increase ecosystem stability; consequently, the response of hunter-gatherer population density to biodiversity seems to saturate as well.

As such, our results indicate that the factors affecting hunter-gatherer abundance on a global scale are similar to those controlling global-scale biodiversity; macroscale studies on biodiversity indicate that, in high-latitude regions, productivity is limiting diversity (30–32), whereas in the tropics, biotic interactions become a more important driver (33). In the case of hunter-gatherers, abundance and stability of resources limit population density in the low-productivity environments, whereas species interactions,
as represented in this study by pathogen stress, become more important in controlling population density in the high-productivity environments. Similarities in the patterns of limiting factors of hunter-gatherer population density and global biodiversity might suggest similarities in the underlying mechanisms as well.

However, because we do not have direct measures of pathogen prevalence among individual study populations, we cannot completely rule out spurious relationships between pathogen stress and hunter-gatherer population density. Due to generally low population densities, it has been suggested that the prevalence of pathogens, such as influenza and measles viruses, and Plasmodium and Trypanosoma parasites is low among hunter-gatherers (34). In addition, it has been argued that high rates of mobility prevent the accumulation of fecal and other waste at hunter-gatherer dwelling sites, leading to a low prevalence of soil-transmitted helminths (35, 36). Even if hunter-gatherers are less susceptible to pathogens than farmers who live under similar conditions of environmental pathogen stress, our findings nevertheless indicate that the interaction between pathogens and hunter-gatherers is sufficient to create a significant negative effect on hunter-gatherer abundance.

By suggesting that subtropical and temperate biomes can sustain the highest hunter-gatherer population densities, our results corroborate the earlier untested hypothesis of Binford (10) regarding the role of resources and pathogens in creating optimal conditions for hunter-gatherers in temperate and subtropical biomes. Based on our results, it seems that, in subtropical and temperate biomes, there is an optimum tradeoff between the opposing effects of resource availability and pathogens: at the higher-productivity end, negative biotic interactions are too limiting, whereas at the lower-productive end, resources are too limiting. However, a large fraction of the area of these highly suitable biomes has been occupied by agricultural populations since at least the late mid-Holocene. It has been argued that stable environments with high resource abundance and richness provide favorable conditions for the innovation and adoption of agriculture (37, 38). Thus, it is understandable that the most suitable environments for hunter-gatherers are also among the first ones to support agriculture. Furthermore, potentially high hunter-gatherer population densities in these areas might have facilitated the innovation, spread, and development of the new subsistence technology. Alternatively, high hunter-gatherer population density itself, resulting in population packing, and reduced options for mobility might have forced hunter-gatherer populations toward agriculture (10, 39). Regardless of the favored explanation, by showing that hunter-gatherer population density and the abundance and stability of their resources are strongly linked, our results highlight the difficulty in distinguishing between environmental and demographic causes of the subsistence change. Given hunter-gatherers’ well-documented ability to affect species communities (40), we can nevertheless assume that the higher the hunter-gatherer abundance that a given biome can potentially sustain, the stronger the impact of anthropogenic drivers on the biome has been even before the transition to agriculture.

In summary, our analyses revealed that NPP, biodiversity, and environmental pathogen stress interact to impose complex and varying limitations on hunter-gatherer population density in different parts of the world. These findings highlight the key role that environmental factors play in shaping the global population density patterns of preagricultural humans. Although cultural
evolution has made humans enormously potent ecosystem engineers and has enabled humans to survive under a variety of conditions (for example, by extracting energy in low-productivity environments and compiling knowledge on medical plants to fight against pathogens), it has not freed hunter-gatherer populations from biotic and abiotic forcing. We can, therefore, assume that spatial and temporal variability in this forcing has been important for our biological and cultural evolution as well.

Materials and Methods

Data. All data manipulation and analyses were conducted using R statistical software (41). The R script and the datasets are freely available at the Zenodo repository (https://doi.org/10.5281/zenodo.1069786), which makes it possible to reproduce this study. SI Materials and Methods gives more details regarding datasets and methods.

We extracted the annual mean temperature and annual precipitation values from the WorldClim.org database (42), and based on these values, we calculated climatic NPP using the empirical Miami model (43) (Fig. S1A). The Miami model, although relatively simple, has been shown to give reasonable estimates of NPP in current climatic and vegetation conditions (44) (SI Materials and Methods). Instead of using satellite-based values, we decided to use climate-based NPP estimates to avoid the potential problems caused by results of NPP and to reach an estimate that reflects more natural levels of NPP.

We used combined mammal, bird, and vascular plant richness as a surrogate of overall biodiversity. Global mammal and bird richness data were extracted from the BiodiversityMapping.org database (45). Global native vascular plant richness data are model based, because native plant species ranges are insufficiently documented in most regions of the world (46, 47).

We interpolated plant richness data to cover the same global grid with the same spatial resolution as the NPP and mammal and bird richness data using inverse distance weighting interpolation with the “gstat” package in R (48–50) (SI Materials and Methods and Fig. S5 A and B). The biodiversity variable was calculated as an average of the scaled animal (mammal + bird) and plant richness (Fig. S1B).

The estimations of environmental pathogen stress are based on observations of environmental conditions prevailing among 186 industrial societies around the globe (51–53). These societies form the Standard Cross-Cultural Sample (www.worldcultures.org) (54). We used the pathogen prevalence index, which is an average of the z scores of the prevalence of 10 pathogens (malaria, dengue, filariae, typhus, trypanosomiasis, leishmaniasis, schistosomes, plague, leprosy, and spirochetes). We interpolated the pathogen data to cover the same global grid with the same spatial resolution as the NPP and mammal and bird richness data using inverse distance weighting interpolation with the “gstat” package in R (48–50) (SI Materials and Methods and Figs. S1C and S5 C and D). We assume that the global interpolation of the pathogen data reflects the relative pathogen stress levels in the local environments of the study populations.

The hunter-gatherer population density data were combined from two ethnographic datasets compiled by Binford (10) and Kelly (11), with the addition of Saami in northern Europe from another source (55). To reduce the uncertainty related to ethnographic population density estimates, if a population was present in both datasets, we averaged the population density values from both sources. When a hunter-gatherer population was present in only one of the datasets, we used the value given by that dataset (SI Materials and Methods and Fig. S6). For each population in the data, we extracted the values of the environmental explanatory variables from the global environmental layers of NPP, biodiversity, and environmental pathogen stress.


Methods. Visual inspections of the relationships between NPP and the log population density of hunter-gatherers and between NPP and environmental pathogen stress indicated shifts in the effects of NPP. This impression was confirmed using breakpoint analysis with the segmented package in R, which allows for the fitting of piecewise linear regression models and provides estimates of breakpoints (56, 57). This package uses an iterative model-fitting algorithm. We did not provide any starting values for the breakpoint parameters, which means that the algorithm used the median of the predictor variable as a starting value for breakpoint estimation. The Davies test (58) included in the segmented package was performed to test whether the difference between the slopes before and after the estimated breakpoint was different from zero. We calculated the mean of the two NPP breakpoints (for population density and pathogen stress) and used this value to distinguish between low- and high-productivity environments.

Structural equation modeling is an approach that interprets information about the observed correlations between system components to evaluate potentially complex causal relationships (23). The approach begins with hypothesizing the underlying structure of causal pathways in the system (a conceptual model or metamodel). Then, this structure is translated into regression equations, and finally, these equations are evaluated against data to support or refute the hypothesized structure. Through this process, structural equation modeling provides an understanding of the nature and magnitude of direct and indirect effects in a system (23), which in our case, is composed of three environmental factors and human population density.

Here, we performed structural equation modeling as a piecewise estimation of local relationships, in which the parameters are estimated separately for each regression equation (59). The fit of the structure of the piecewise model is evaluated using a test of directed separation, with Fisher’s C as the test statistic (59, 60). This test indicates whether missing paths between variables in a system should be included in the model. We performed modeling using the R package piecewiseSEM, which allows for the automating of fitting of lists of structured equations and goodness-of-fit tests (59). For the sake of interpretability, we fit separate models for low- and high-productivity environments.

We began with the models (for low- and high-productivity environments) that omitted the direct effects of NPP on log, population density. For low- and high-productivity structural equation models, the goodness-of-fit test strongly indicated that the direct effect of NPP should be included. In such saturated models, however, the effect of pathogen stress on population density (low productivity) and the effect of biodiversity on both pathogen stress and population density (high productivity) were nonsignificant. After omitting these effects, the goodness-of-fit test indicated that the structural equation models fit the data well in both low- and high-productivity environments.

However, after adjusting the P values for spatial autocorrelation (61) (SI Materials and Methods), the effect of NPP on pathogen stress turned out to not be statistically significant in the low-productivity environments. Therefore, we omitted this final effect. Although Fisher’s C indicated a lower fit than for the model that included the effect of NPP on pathogen stress.

For global projections of hunter-gatherer population density, we used two regression methods: ordinary least squares linear regression model (LMI) (Fig. S3 and Table S2) and support vector regression (SVR) (62–64). Both of these models included interaction terms between biodiversity and NPP and between pathogen stress and NPP and SVR models were trained to reproduce high-resolution predictions (Fig. S7). However, because fitting algorithms are different (62–64), their predictions are not identical, and therefore, these techniques are complementary (Fig. S7). Neither of the techniques seem to extrapolate, and the predicted values are within the boundaries of the original hunter-gatherer data (Fig. S7). In the context of global prediction and extrapolation, we prefer to use a combination of two techniques instead of single technique.

We fit global models using NPP, biodiversity, and pathogen stress as predictors of log, population density. SVR was implemented with the e1071 package in R (65). The LM and SVR (with linear kernel) models were trained without splitting the ethnographic data into low- and high-productivity subsets. After the training, we used model algorithms to predict the log, population density using global environmental data. We calculated the mean of the predictions of the LM and SVR models and used this average as our global projection of hunter-gatherer population densities as shown in Fig. 3B. The performance of the averaged (LM and SVR) model was evaluated using h-block cross-validation (SI Materials and Methods) (66, 67). h-block cross-validated (h = 500 km) prediction root mean squared error is 1.26, and R² is 0.46. These metrics indicate that the model is performing well with the unseen data.

The limiting effect of a predictor variable is calculated as the difference between the population density of the prediction, in which the given predictor is set to its optimum value, and the “true” prediction (see above), in which all of the predictor variables are allowed to vary. The larger the difference, the stronger the limiting effect of the predictor. The map of the limiting effects of NPP, biodiversity, and pathogen stress was constructed by determining the optimum values of each predictor in the averaged (LM and SVR) model. The optimum value of a predictor is the value where the log, population density reaches its highest value in the ethnographic data, as seen in the response shapes of the averaged (LM and SVR) model. We then created three global datasets, in which one predictor variable at a time was

Tallavaara et al.  
PNAS Early Edition | 5 of 6

[Image]
density data. This allowed us to extract information about the distributions of the limiting effects and the log, population density in different biomes. For the biowise comparisons, the limiting effects were scaled to sum to unity.

**ACKNOWLEDGMENTS.** We thank M. Fortelius, J. Kankaanpää, R. L. Kelly, T. Rankama, and three anonymous reviewers for their constructive critique. M.T. and J.T.E. acknowledge financial support from the Kone Foundation. J.T.E. acknowledges this as a contribution to the Integrative Climate Change Biogeography Programme under International Union of Biological Sciences.

7. Quat Sci Rev 93:2241
21. Hochachka WM, Dhondt AA (2000) Density-dependent decline of host abundance and species diversity in these regions is further away from its optimum value than it would be without the recent human influence.

For biowise comparisons, we rasterized the biome dataset (see above) and aligned it with the global projected limiting effects and the population density data. This allowed us to extract information about the distributions of the limiting effects and the log, population density in different biomes. For the biowise comparisons, the limiting effects were scaled to sum to unity.