

UC Santa Barbara

UC Santa Barbara Previously Published Works

Title

Quantitative historical analysis uncovers a single dimension of complexity that structures global variation in human social organization

Permalink

<https://escholarship.org/uc/item/7x48671z>

Journal

Proceedings of the National Academy of Sciences of the United States of America, 115(2)

ISSN

0027-8424

Authors

Turchin, Peter
Currie, Thomas E
Whitehouse, Harvey
et al.

Publication Date

2018-01-09

DOI

10.1073/pnas.1708800115

Peer reviewed



Quantitative historical analysis uncovers a single dimension of complexity that structures global variation in human social organization

Peter Turchin^{a,b}, Thomas E. Currie^{c,1}, Harvey Whitehouse^{d,e}, Pieter François^{d,f}, Kevin Feeney^g, Daniel Mullins^{d,h}, Daniel Hoyerⁱ, Christina Collins^c, Stephanie Grohmann^d, Patrick Savage^d, Gavin Mendel-Gleason^g, Edward Turnerⁱ, Agathe Dupeyronⁱ, Enrico Cioniⁱ, Jenny Reddishⁱ, Jill Levineⁱ, Greine Jordanⁱ, Eva Brandl^{i,j}, Alice Williams^c, Rudolf Cesaretti^k, Marta Krueger^l, Alessandro Ceccarelli^m, Joe Figliulo-Rosswurmⁿ, Po-Ju Tuanⁱ, Peter Peregrine^{o,p}, Arkadiusz Marciniak^l, Johannes Preiser-Kapeller^q, Nikolay Kradin^r, Andrey Korotayev^s, Alessio Palmisano^t, David Baker^u, Juley Bidmead^v, Peter Bol^w, David Christian^u, Connie Cook^{x,y}, Alan Covey^z, Gary Feinman^{aa}, Ární Daniel Júlíusson^{bb}, Axel Kristinsson^{cc}, John Miksic^{cd}, Ruth Mostern^{ee}, Cameron Petrie^{mm,ff}, Peter Rudiak-Gould^{gg}, Barend ter Haar^{hh}, Vesna Wallace^v, Victor Mairⁱⁱ, Liye Xie^{jj}, John Baines^{kk}, Elizabeth Bridges^{ll}, Joseph Manning^{mm}, Bruce Lockhartⁿⁿ, Amy Bogaard^{oo}, and Charles Spencer^{pp,1}

^aDepartment of Ecology and Evolutionary Biology, University of Connecticut, CT 06269; ^bComplexity Science Hub Vienna, 1080 Wien, Austria; ^cHuman Behaviour & Cultural Evolution Group, Department of Biosciences, University of Exeter, Cornwall TR10 9FE, United Kingdom; ^dInstitute of Cognitive and Evolutionary Anthropology, University of Oxford, Oxford OX2 6PE, United Kingdom; ^eMagdalen College, Oxford OX1 4AU, United Kingdom; ^fSt. Benet's Hall, Oxford OX1 3LN, United Kingdom; ^gSchool of Computer Science and Statistics, Trinity College Dublin, Dublin 2, Ireland; ^hInstitute of English Studies, University of London, London WC1E 7HU, United Kingdom; ⁱSeshat: Global History Databank, Evolution Institute, San Antonio, FL 33576; ^jDepartment of Anthropology, University College London, London WC1H 0BW, United Kingdom; ^kSchool of Human Evolution and Social Change, Arizona State University, Tempe, AZ 85287; ^lInstitute of Archaeology, Adam Mickiewicz University, 61-614 Poznań, Poland; ^mDepartment of Archaeology, University of Cambridge, Cambridge CB2 3DZ, United Kingdom; ⁿDepartment of History, University of California Santa Barbara, Santa Barbara, CA 93106; ^oAnthropology and Museum Studies, Lawrence University, Appleton, WI 54911; ^pSanta Fe Institute, Santa Fe, NM 87501; ^qDivision for Byzantine Research, Institute for Medieval Research, Austrian Academy of Sciences, 1020 Wien, Austria; ^rDepartment of Anthropology, Institute of History, Archaeology and Ethnology, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok 690001, Russia; ^sLaboratory of Monitoring of Destabilization Risks, National Research University Higher School of Economics, Moscow 125267, Russia; ^tInstitute of Archaeology, University College London, London WC1H 0PY, United Kingdom; ^uBig History Institute, Macquarie University, Sydney NSW 2109, Australia; ^vDepartment of Religious Studies, Chapman University, Orange, CA 92866; ^wEast Asian Languages and Civilizations, Harvard University, Cambridge, MA 02138; ^xDepartment of Modern Languages & Literatures, Lehigh University, Bethlehem, PA 18015; ^ySchool of Historical Studies, Institute of Advanced Studies, Princeton, NJ 08540; ^zDepartment of Anthropology, University of Texas Austin, Austin, TX 78712; ^{aa}Integrative Research Center, Field Museum of Natural History, Chicago, IL 60605; ^{bb}History, University of Iceland, IS-108 Reykjavík, Iceland; ^{cc}History, Reykjavík Academy, IS-108 Reykjavík, Iceland; ^{cd}Department of Southeast Asian Studies, National University of Singapore, Singapore 119260; ^{de}Department of History, University of Pittsburgh, Pittsburgh, PA 15260; ^{ff}Trinity College, Cambridge CB2 1TQ, United Kingdom; ^{gg}Independent Scholar, Toronto, ON M6P 1T6, Canada; ^{hh}Faculty of Oriental Studies, University of Oxford, Oxford OX2 6LU, United Kingdom; ⁱⁱDepartment of East Asian Languages and Civilizations, University of Pennsylvania, Philadelphia, PA 19104; ^{jj}Department of Anthropology, University of Toronto, Toronto, ON M5S 2S2, Canada; ^{kk}Faculty of Oriental Studies, Oriental Institute, University of Oxford, Oxford OX1 2LE, United Kingdom; ^{ll}Institute of Archaeology and Anthropology, University of South Carolina, Columbia, SC 29208; ^{mm}Department of History, Yale University, New Haven, CT 06520; ⁿⁿDepartment of History, National University of Singapore, Singapore 117570; ^{oo}Institute of Archaeology, University of Oxford, Oxford OX1 2PG, United Kingdom; and ^{pp}Division of Anthropology, American Museum of Natural History, New York, NY 10024

Contributed by Charles Spencer, November 16, 2017 (sent for review May 26, 2017; reviewed by Simon A. Levin and Charles Stanish)

Do human societies from around the world exhibit similarities in the way that they are structured, and show commonalities in the ways that they have evolved? These are long-standing questions that have proven difficult to answer. To test between competing hypotheses, we constructed a massive repository of historical and archaeological information known as “Seshat: Global History Databank.” We systematically coded data on 414 societies from 30 regions around the world spanning the last 10,000 years. We were able to capture information on 51 variables reflecting nine characteristics of human societies, such as social scale, economy, features of governance, and information systems. Our analyses revealed that these different characteristics show strong relationships with each other and that a single principal component captures around three-quarters of the observed variation. Furthermore, we found that different characteristics of social complexity are highly predictable across different world regions. These results suggest that key aspects of social organization are functionally related and do indeed coevolve in predictable ways. Our findings highlight the power of the sciences and humanities working together to rigorously test hypotheses about general rules that may have shaped human history.

cultural evolution | sociopolitical complexity | comparative history | comparative archaeology | quantitative history

The scale and organization of human societies changed dramatically over the last 10,000 y: from small egalitarian groups integrated by face-to-face interactions to much larger societies with specialized governance, complex economies, and sophisticated

information systems. This change is reflected materially in public buildings and monuments, agricultural and transport infrastructure, and written records and texts. Social complexity, however, is a characteristic that has proven difficult to conceptualize and quantify (1, 2). One argument is that these features of societies are functionally interrelated and tend to coevolve together in predictable ways (3, 4). Thus, societies in different places and at different points in time can be meaningfully compared using an overall

Author contributions: P.T., T.E.C., H.W., P.F., and K.F. designed research; P.T., T.E.C., H.W., P.F., K.F., D.M., D.H., C. Collins, S.G., G.M.-G., E.T., A.D., E.C., J.R., J.L., G.J., E. Brandl, A.W., R.C., M.K., A. Ceccarelli, J.F.-R., P.P., and A.P. performed research; P.T., T.E.C., and P.S. analyzed data; D.M., D.H., C. Collins, S.G., and G.M.-G. participated in the conceptual development of data coding schemes and supervised data collection; E.T., A.D., E.C., J.R., J.L., G.J., E. Brandl, A.W., R.C., M.K., A. Ceccarelli, J.F.-R., and P.-J.T. collected the data and contributed to the development of data coding schemes; P.P., A.M., J.P.-K., N.K., A. Korotayev, A.P., D.B., J. Bidmead, P.B., D.C., C. Cook, G.F., Á.D.J., A. Kristinsson, J.M., R.M., C.P., P.R.-G., B.t.H., V.W., V.M., L.X., J. Baines, E. Bridges, J. Manning, B.L., A.B., and C.S. guided data collection, checked data for their domains of expertise, and contributed to the conceptual development of data coding schemes; and P.T., T.E.C., and C.S. wrote the paper.

Reviewers: S.A.L., Princeton University; and C.S., University of California, Los Angeles.

The authors declare no conflict of interest.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Data deposition: We have created a publicly accessible website (seshatdatabank.info) that shows how entries in “Seshat: Global History Databank,” are supported by references, and explanations and justifications of the codes.

¹To whom correspondence may be addressed. Email: T.Currie@exeter.ac.uk or cspencer@amnh.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1708800115/-DCSupplemental.

Significance

Do human societies from around the world exhibit similarities in the way that they are structured and show commonalities in the ways that they have evolved? To address these long-standing questions, we constructed a database of historical and archaeological information from 30 regions around the world over the last 10,000 years. Our analyses revealed that characteristics, such as social scale, economy, features of governance, and information systems, show strong evolutionary relationships with each other and that complexity of a society across different world regions can be meaningfully measured using a single principal component of variation. Our findings highlight the power of the sciences and humanities working together to rigorously test hypotheses about general rules that may have shaped human history.

measure of social complexity (2). Several researchers have attempted to come up with a single measure to capture social complexity (5–7), but a more common approach has been to use proxy measures, such as the population size of the largest settlement (7, 8), number of decision-making levels (9), number of levels of settlement hierarchy (10), or extent of controlled territory (11). Others have criticized this approach on the grounds that these proposed measures focus too narrowly on size and hierarchy (12, 13) or that there are multiple dimensions or variable manifestations of complexity (14). However, another common view is that different societies have unique histories and cannot be meaningfully compared in this way (15). Indeed, most historians have abandoned the search for general principles governing the evolution of human societies (16, 17). However, although every society is unique in its own ways, this does not preclude the possibility that common features are independently shared by multiple societies. How can we study both the diversity and commonalities in social arrangements found in the human past?

In this paper, we address these issues by building a global historical and archaeological database that takes into account the fragmentary and disputed nature of information about the human past. To test hypotheses about the underlying structure of variation in human social organization, we apply a suite of statistical techniques to these data, including principal component analysis (PCA). We then compare evolutionary trajectories in world regions by plotting the estimated first principal component (PC) of variation against time.

Building a Comparative Database of Human History

Previous attempts to address these questions have been limited by a reliance on verbal arguments (15, 18, 19), comparisons involving a small number of polities (20, 21), noncomprehensive data samples (3, 22), or nonsystematic methods of data coding and purely descriptive analyses (6, 23–25). To advance beyond purely theoretical debates and comparisons based on limited samples, we have built a massive repository of systematically collected, structured historical and archaeological data known as “Seshat: Global History Databank” (26) (*Materials and Methods*). In collecting data, we used a targeted, stratified sampling technique that aims to maximize the variation in forms of social organization captured from as wide a geographic range as possible [thus minimizing pseudoreplication of data points (27)]. Specifically, we divided the world into 10 regions and in each, selected three locations or “Natural Geographic Areas” (NGAs), representing early, intermediate, and late appearance of politically centralized societies (Fig. 1). The construction of this databank has been accomplished in collaboration with a large number of historical and archaeological experts. Our goal is to capture the state of the art knowledge about past societies, including where information

is uncertain or there are disagreements between researchers (*Materials and Methods*). The online version of the databank (seshatdatabank.info/) illustrates how entries in the databank are supported by explanations of coding choices and references (*SI Appendix, SI Methods*).

Our unit of analysis is a polity: an independent political unit that ranges in scale from groups organized as independent local communities to territorially expansive, multiethnic empires. To populate the databank, we coded information on all identifiable polities ($n = 414$) that occupied each of the 30 NGAs at 100-y time slices from the beginnings of agriculture (in some cases, as far back as 9600 BCE) to the modern period (in some cases, as late as 1900 CE) (*SI Appendix, SI Methods*). To capture different aspects of social complexity, we systematically collected data on 51 variables that could be reliably identified and categorized from the historical and archaeological records. These variables were then aggregated into nine “complexity characteristics” (CCs) (Fig. 24). The first set of variables relates to the size of polities: polity population (CC1), extent of polity territory (CC2), and “capital” population (the size of the largest urban center; CC3). A second set of variables measures hierarchical complexity (CC4), focusing on the number of control/decision levels in the administrative, religious, and military hierarchies and on the hierarchy of settlement types (village, town, provincial capital, etc.). Government (CC5) variables code for the presence or absence of official specialized positions that perform various functions in the polity: professional soldiers, officers, priests, bureaucrats, and judges. This class also includes characteristics of the bureaucracy (e.g., presence of an examination system), the judicial system, and specialized buildings (e.g., courts). Infrastructure (CC6) captures the variety of observable structures and facilities that are involved in the functioning of the polity. Information system (CC7) codes the characteristics of writing, record-keeping, etc. We also record whether the society created literature on specialized topics, including history, philosophy, and fiction (texts; CC8). Finally, economic development is reflected in monetary system (CC9), which represents the “most sophisticated” monetary instrument present in the coded society, and indicates the degree of economic complexity that would be possible. Our data collection process also allows us to incorporate uncertainty in this coding or disagreement among sources (*Materials and Methods*).

Testing Hypotheses About the Evolution of Social Complexity

To test between the different hypotheses laid out above, we analyzed these data using PCA, which assesses the extent to which different variables are tapping into shared dimensions of variation. We expected CC1–CC3 to cluster tightly together, as they all measure size, albeit in somewhat different ways. Beyond this, if the variation in social organization across different societies can be meaningfully captured by a single measure of social complexity, we would predict that the different CCs would correlate strongly with each other and be captured in one PC of variation onto which all CCs load. If social complexity is predictably multidimensional, then other PCs capturing significant amounts of variation might also be present.

We hypothesized that social complexity could be captured by two PCs (7). Size variables (CC1–CC3) should exhibit a strong relationship with hierarchical organization (CC4), as hierarchy is often thought to be a necessary mechanism for enabling effective information flows in large polities (19). We refer to the combination of size and hierarchy as “scale” (Fig. 24). The other variables might form another dimension of “nonscale” complexity, perhaps reflecting specialization of roles and the products that emerge from such specialization. Another possibility is that these CCs covary in other ways or are free to vary independently (that is, they do not evolve together in a predictable manner). In the latter situation, we would not expect correlational analysis or the PCA to reveal any structure in terms of the relationships of these variables with each other.

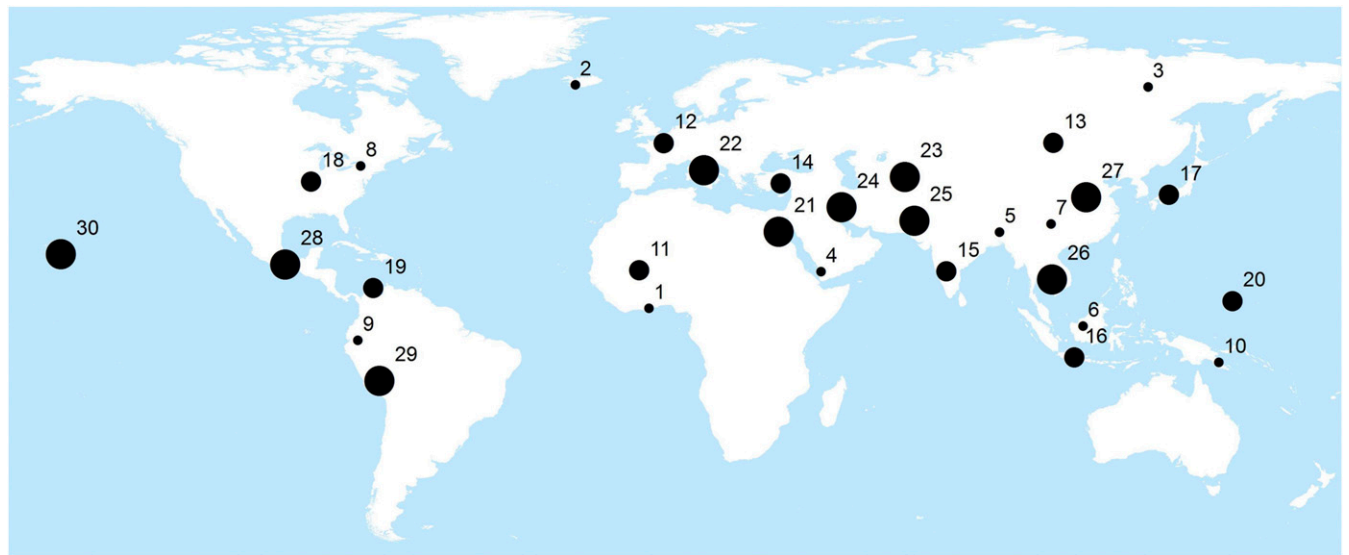


Fig. 1. Locations of the 30 sampling points on the world map (the size of the dot reflects the antiquity of centralized societies within the world region). The key to the numbers is in *SI Appendix, Table S1*.

Contrary to these expectations, all nine CCs showed substantial and statistically significant correlations with each other, with coefficients ranging from 0.49 to 0.88 (*SI Appendix, Table S4*). We found that a single PC, PC1, explains $77.2 \pm 0.4\%$ of variance. The proportion of variance explained by other PCs drops rapidly toward zero (Fig. 2*B*). Furthermore, all CCs load equally strongly onto PC1, indicating that PC1 captures contributions from across the multiple measures of social organization used here (Fig. 2*C* and *SI*

Appendix). This result provides strong support for the hypothesis that social complexity can be captured well by a single measure. In running these analyses, we have to take into account a number of factors, including missing data and various sources of autocorrelation. However, our results are robust to a large number of different assumptions and potential sources of error and bias (*SI Appendix, SI Results*).

We can also test directly the idea that societies that developed on distant world continents share enough similarities in their

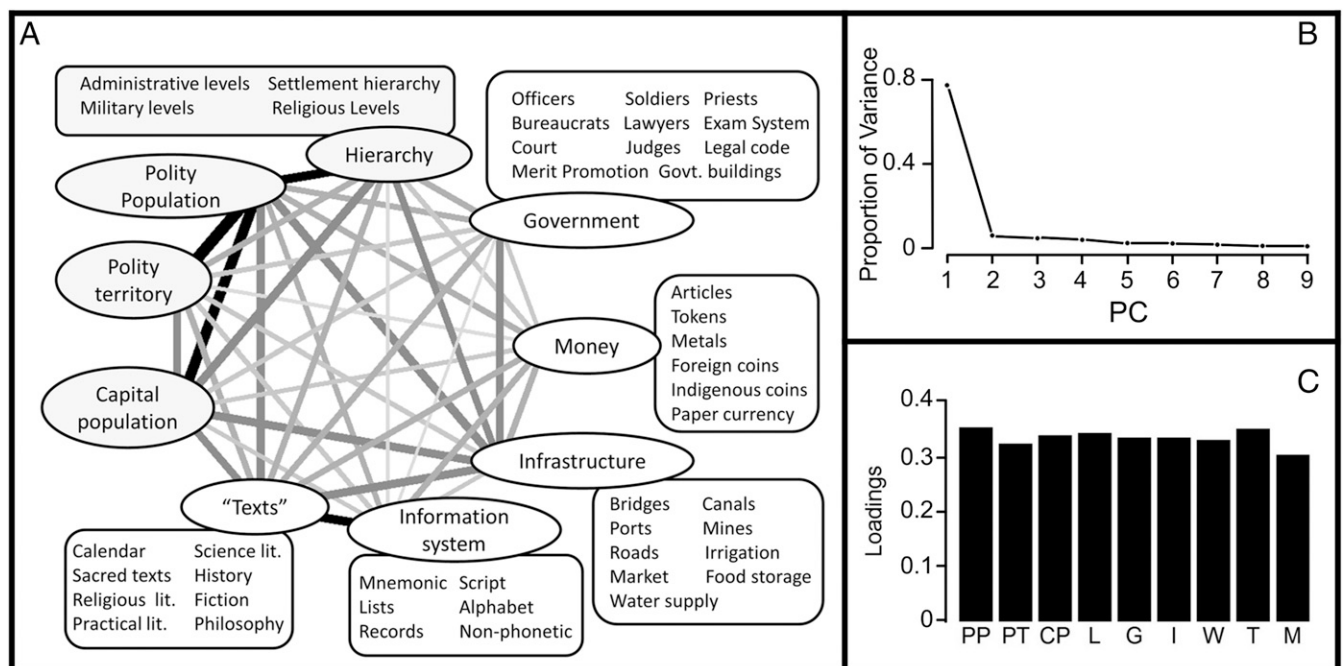


Fig. 2. (A) Nine CCs (ovals) aggregating 51 variables (*SI Appendix* has details on all CCs). Line width and color are proportional to the correlation coefficients between CCs (darker and thicker lines indicate stronger correlations). All CCs are significantly correlated with one another (correlation coefficients range between 0.49 and 0.88). Some variables show stronger linkages with each other, such as the scale variables (ovals shaded in gray), whereas money is less strongly correlated with the other variables. (B) Proportion of variance explained by PCs. (C) Factor loadings for CCs on PC1 indicating strong contributions by all CCs to a single dimension of social complexity. CP, capital population; G, government; I, infrastructure; L, levels; M, money; PP, polity population; PT, polity territory; T, texts; W, information system (writing).

Table 1. Cross-validation results for out of sample prediction of CCs across all world regions

Predicted CC	Overall ρ^2
Polity population	0.84
Polity territory	0.76
Capital population	0.71
Levels of hierarchy	0.60
Government	0.53
Infrastructure	0.62
Information system	0.59
Texts	0.73
Monetary system	0.53

Prediction accuracy is measured with prediction ρ^2 (*SI Appendix, Table S2*). Overall ρ^2 values are calculated as an average of the ρ^2 values weighted by the number of polities from which they are drawn.

complexity dimensions to allow for meaningful comparisons. We used the statistical technique of k -fold cross-validation (28), in which models are fitted on one set of data (“training set”) and evaluated on another independent set (“testing set”). We reserved all data for polities in a particular world region, such as North America, as the testing set; developed predictive models on the rest of the data (by regressing each CC in turn on other CCs); and then, used the fitted models to predict each CC for North American polities. We then repeated this analysis for all other world regions. The accuracy of prediction is measured by the coefficient of prediction, ρ^2 , which approaches one if prediction is very accurate, takes the value of zero when prediction is only as good as simply using the mean, and can take negative values if model prediction is worse than the mean.

Our results show that the values of CCs can be predicted by knowledge of other CCs (Table 1), and as Table 2 shows, median ρ^2 ranges between 0.08 (Southeast Asia) and 0.91 (North America), indicating that this predictive ability holds across all world regions. Low ρ^2 values do occur for some variables and seem to be lowest for those regions with the fewest number of polities to be predicted (*SI Appendix, SI Results*). This is to be expected, as with fewer cases to predict, there is less chance for general relationships to be detected. Some decreases in ρ^2 may also occur if smaller societies adopt some of the features, which make up CCs, from other societies, because they may be useful in dealing with larger societies (perhaps especially aspects of money and writing). Such selective adoption may not necessarily lead to the rapid devel-

Table 2. Cross-validation results for out of sample prediction of CCs summarized for different world regions

Predicted region	ρ^2			n
	Median	Minimum	Maximum	
Africa	0.72	0.37	0.90	41
Central Eurasia	0.63	−0.38	0.86	9
East Asia	0.70	0.30	0.93	34
Europe	0.53	−0.31	0.84	43
North America	0.91	0.79	0.97	11
Oceania–Australia	0.14	−3.21	0.97	1
South America	0.74	−24.57	0.97	5
South Asia	0.46	−0.05	0.69	12
Southeast Asia	0.08	−4.27	0.91	8
Southwest Asia	0.71	0.19	0.79	39
All regions	0.62	0.53	0.84	203

Prediction accuracy is measured with prediction ρ^2 (*SI Appendix, Table S2*). Median, minimum, and maximum indicate the median, smallest, and largest ρ^2 values across the nine CCs for the region, respectively.

opment of other aspects of complexity. Lower ρ^2 may also occur if some traits are retained when others are lost (see below).

Comparing Evolutionary Trajectories

Our results, thus, indicate that there is striking similarity in the way that the societies in our global historical sample are organized. Examining PC1 enables us to compare how social complexity evolved in different parts of the globe over time. We plotted PC1 values estimated for each polity that occupied each of the 30 NGAs at 100-y time intervals. Fig. 3 compares the trajectories of the NGAs with early appearance of politically centralized societies in each of the 10 world regions (*SI Appendix* has all 30 trajectories). These trajectories indicate a general increase in complexity over time, albeit with occasionally substantial decreases in complexity (29). This comparison shows that there are crucial differences in the timing of takeoff and the rate of change as well as level of social complexity reached in different regions by 1900—differences that become clearly revealed through the analyses performed here. For example, although it is well-known that complex societies of the Americas emerged later than those in Eurasia, using our data, we can quantify their differences in social complexity. The difference in PC1 levels indicates that societies in the Americas were not as complex as those from Eurasia at time of contact, which may be a contributing factor in explaining why European societies were able to invade and colonize the Americas (30).

The tight relationships between different CCs provide support for the idea that there are functional relationships between these characteristics that cause them to coevolve (3). Scale variables are likely to be tightly linked, since increases or decreases in size may require changes in the degree of hierarchy (both too few and too many decision-making levels create organizational problems) (19). A similar argument has been put forward for size and governance (20). The production of public goods, such as infrastructure, may require solutions to collective action problems (31), and these can be provided by governance institutions and professional officials (32). Despite these linkages, because of their nature, different CCs are likely to show different temporal dynamics. Levels of nonscale characteristics, such as information systems, monetary systems, or infrastructure, may be retained and used even if a polity does decrease in size. Indeed, by retaining such features, the scale of the polity may more readily bounce back and return to its former level. This cultural continuity may be one reason why the trends that we see in our data are for social complexity to increase over time in a cumulative, ratchet-like manner (3, 33–35). For example, polities in our Italian NGA had writing, texts, and coins before the dramatic rises in scale of the Roman republic and empire, and they retained these features after the fall of Rome.

Discussion

One major conclusion from these analyses is that key aspects of human social organization tend to coevolve in predictable ways. This result supports the hypothesis that there are substantial commonalities in the ways that human societies evolve. Thus, societies can be meaningfully compared along a single dimension, which can be referred to as social complexity. Our analyses suggest that the estimated first PC of social complexity can be interpreted as a composite measure of the various roles, institutions, and technologies that enable the coordination of large numbers of people to act in a politically unified manner. However, as noted in the Introduction to this paper, the term “social complexity” has previously been defined and discussed in many ways. Indeed, complexity is a term that has many colloquial meanings, and there are many valid ways in which it could be applied to human social organization. For example, the kinship systems of some Australian Aboriginal groups, such as the Aranda, involve many complicated rules that determine who can marry whom (36, 37), and Turkana pastoralists have sophisticated social rules

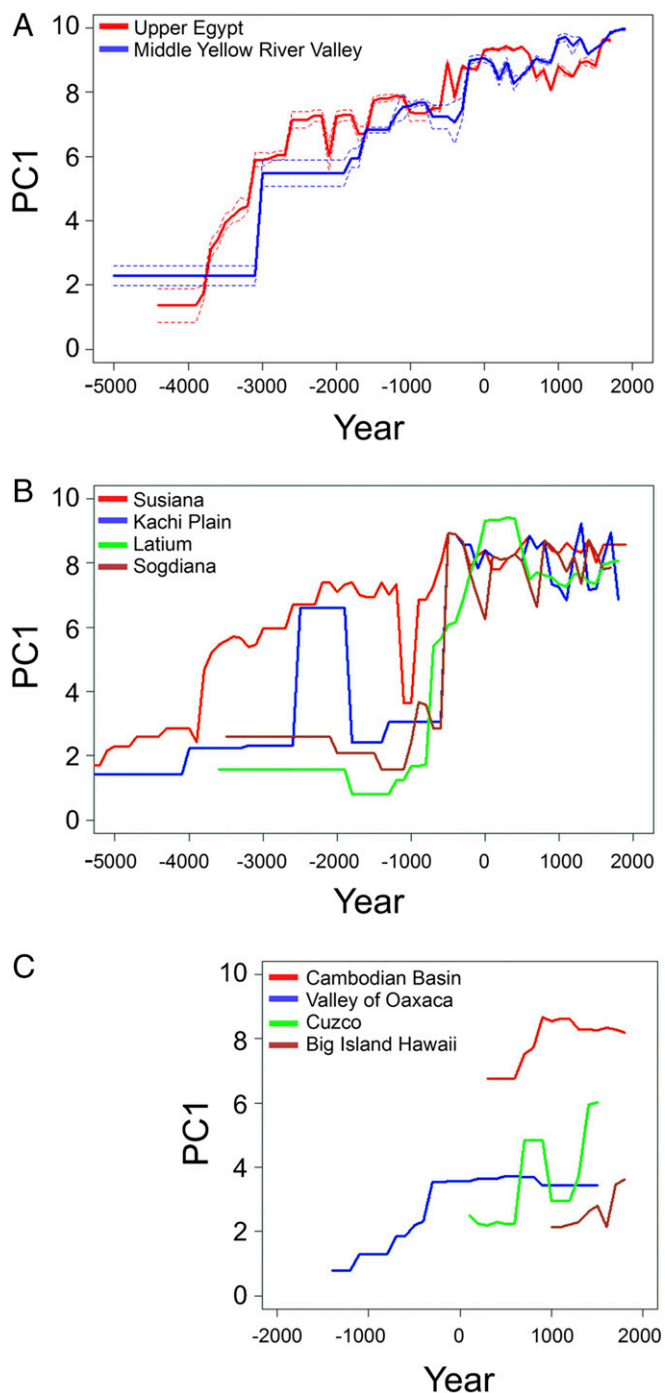


Fig. 3. Trajectories of social complexity in 10 world regions quantified by PC1 values for locations where centralized, hierarchical polities first appeared in a particular region. (A) Africa and east Asia. Broken lines indicate 95% confidence intervals. (B) Southwest Asia, south Asia, Europe, and central Asia. (C) Southeast Asia, North America, South America, and Oceania. Confidence intervals for B and C are shown in *SI Appendix, Figs. S4 and S5*. PC1 has been rescaled to fall between 0 (low complexity) and 10 (high complexity) to aid interpretation. Flat horizontal lines indicate periods when there is no evidence of change from our polity data.

and norms that enable them to join together in large groups to conduct cooperative raiding missions (38).

Building historical databases, such as Seshat, allows us to take the vast amount of information about the human past and use it to test and reject competing hypotheses in the same cumulative process that characterizes the sciences (39, 40). It is important to

emphasize that we attach no normative judgment to the measure of social complexity that we have identified here; more complex societies are not necessarily “better” than less complex societies. We need to separate out these issues as well as ethnocentric judgments about non-European societies (2) from the kind of questions about how societies have actually evolved that we address here (3).

Our purpose here is not to propose that one definition of social complexity is superior to another. Instead, by supplying evidence that at least some aspects of human societies evolve in predictable and interconnected ways, this study illustrates that it is possible to move beyond the kind of verbal arguments that too often dominate debates about the evolution of human social organization. Furthermore, quantitative comparative analysis forces us to be more explicit about the evidence needed to support different claims and brings greater clarity to debates and discussions. It is important to recognize that, in any study, including this one, there are many subjective judgments about the coding of variables. Our goal in establishing the databank is to provide a summary of what is currently known about past human societies based on the literature and the expert knowledge of academics. It is not our aim to provide a more objective or definitive representation of such evidence but rather, to make the decisions and assumptions behind our data more explicit than has often been the case in the past. Our databank thus allows others viewing these data to challenge these decisions and provide alternative assessments. Future analyses can then assess whether alternative coding decisions substantially affect the results presented here.

The choice of variables and CCs themselves is also an important consideration in evaluating these results. We have attempted to be inclusive by choosing variables that would not favor particular forms of governance from certain parts of the world as being more complex. The variables are broad enough to allow for such features to come from a variety of specific institutions and are not biased toward Western forms of governance, which ultimately have their origins in early states in Greece and Mesopotamia. Our government variables (CC5), for example, capture the degree of specialization and professionalization of those involved in decision-making in sociopolitical affairs, a characteristic that has long been central to discussion of social complexity in different parts of the world (41). Our information system and texts variables (CC7 and CC8, respectively) capture the extent to which different types of information are being recorded and transmitted and reflect diversity and specialization in learning. Such information is potentially important in organizing societies or enabling societies to solve adaptive problems. Again, the variables within this category are broad enough to not be specific to any particular cultural tradition a priori. In particular, writing has been independently invented in such distant world regions as western Eurasia, east Asia, and Mesoamerica. As with the coding of specific variables, future analyses could assess whether the inclusion of alternative variables substantially affects the results presented here. Importantly, if our choice of variables was biased toward certain cultural-historical traditions, then this would reduce the correlations between different aspects of complexity, and these patterns would be different in different parts of the world. However, the overall high degree of correlation between CCs, as our cross-validation results indicate, suggests that the patterns that we have identified are relatively stable across regions.

The approach that we have taken in this paper can be used to resolve other long-standing controversies in the study of human societies. For example, some researchers have argued that traditional approaches to social complexity have overemphasized hierarchical relationships and did not pay enough attention to more horizontal or heterarchical forms of complexity (13, 42). Power relationships within societies can range from being autocratic or exclusionary (certain individuals or groups aim to control sources of power) to more corporate/collective, in which power is broadly shared across different sectors of societies (12, 43, 44). Other authors have

identified additional patterns that might be seen in human social evolution (21, 45), which can be fruitfully studied with the approach in this article. Indeed, some of the features that we have already coded, such as types and numbers of official positions, could be important in addressing such issues. We are already collecting data to test the idea that the balance between autocratic and collective forms of power has changed systematically over time, with autocratic forms being more prevalent in chiefdoms and early states. The emergence of institutions that held despotic leaders to account is argued to have occurred later (26), perhaps in connection with the emergence of certain religions (46, 47).

Our approach is also well-suited to go beyond identifying patterns and investigate the processes of sociopolitical evolution. The systematic compilation of long-term diachronic data for multiple variables on a large number of societies has been relatively rare in comparative history and archaeology (refs. 20, 35, and 48–50 have comparative studies of evolutionary trajectories for a smaller number of cases or time periods). Previous large-scale comparative approaches have generally focused on comparing evolutionary outcomes (end points) or snapshots at a single period of time rather than entire long-term trajectories (25, 51–54). By analyzing trajectories, we can both examine the processes that lead to variation in human societies across space and time and also take into account the historical changes that are contingent on the particular conditions and past history of the societies involved (3, 4, 55, 56).

In this study, the focus on looking at comparative changes over time enables us to investigate questions about the tempo of evolutionary change in human social systems. One pattern that is already apparent (Fig. 3 and *SI Appendix, Fig. S6*) is that many trajectories exhibit long periods of stasis or gradual, slow change interspersed with sudden large increases in the measure of social complexity over a relatively short time span. This pattern is consistent with a punctuational model of social evolution, in which the evolution of larger polities requires a relatively rapid change in sociopolitical organization, including the development of new governing institutions and social roles, to be stable (3, 4, 57). One example that has been investigated in previous work is the emergence of bureaucratic forms of governance, which tend to develop around the time when polities first extend political control beyond more than a day's round trip from the capital (20). A related idea is that, if there are strong relationships between these variables and if change is relatively rapid, then societies may tend to evolve toward certain types of sociopolitical organization based on associations between certain combinations of traits (3, 24, 57). Cluster analysis of PC1 shows some initial support for this idea, indicating a clear distinction between large societies that exhibit many of the nonscale features of complexity and smaller societies that lack most of these features, with other potential groupings within these clusters (*SI Appendix, SI Discussion and Figs. S12 and S13*).

Our data also indicate a shift toward more complex societies over time in a manner that lends support to the idea of a driving force behind the evolution of increasing complexity (3, 10, 58, 59) (*SI Appendix, SI Discussion, Fig. S11, and Table S9*). Such a driven trend is consistent with the hypothesis that competition between groups, particularly in the form of warfare, has been an important selective force in the emergence and spread of large, complex societies (10, 11, 60). In future work, the kind of systematic approach that we have used here will allow us to assess the large number of alternative mechanisms that have been proposed to explain the evolution of social complexity (2, 11, 14, 26). We are currently expanding the Seshat databank to collect information on agricultural productivity, warfare, religion, ritual, institutions, equity, and wellbeing in past societies to assess such competing hypotheses (26, 47, 61, 62).

Our focus in this paper has been on the increase in social complexity over time. However, understanding the causes of collapses

and decreases in social complexity is an equally important research topic. As is clear in the evolutionary trajectories (Fig. 3 and *SI Appendix, Fig. S6*), declines in social complexity, some quite dramatic, are frequently seen in most NGAs. Furthermore, some of the large decreases are “hidden” when a polity collapses, but the NGA is immediately taken over by another large-scale society nearby. While different analytical approaches than the ones used in this article and additional data will be needed to study the processes explaining social collapse, such an investigation is entirely within the scope of the Seshat project.

In summary, our results indicate that it is indeed possible to meaningfully compare the complexity of organization in very different and unconnected societies along a single dimension (6, 30). Although societies in places as distant as Mississippi and China evolved independently on different continents and followed their own trajectories, the structure of social organization, as captured by the interrelations between different CCs, is broadly shared across all continents and historical eras. Key elements of complex social organization have thus coevolved in highly consistent ways across time and space. Differences in the timing of takeoff, the overall rate of increase, and the depth of periodic declines in social complexity provide us with highly informative data for testing theories of social and cultural evolution. Our databank was built via a collaborative relationship with humanities scholars who provided expert knowledge of past societies and helped guide data collection at all stages. This paper has shown the power of the sciences and the humanities working together to help us better understand the past by testing and rejecting alternative hypotheses about the general rules that have shaped human history.

Materials and Methods

Data. Data were collected as part of “Seshat: Global History Databank” (26) (*SI Appendix, SI Methods*). We collected data in a systematic manner by dividing the world into 10 major regions (Fig. 1 and *SI Appendix, Fig. S1 and Table S1*). Within each region, we selected three NGAs to act as our basic geographical sampling unit. Each NGA is spatially defined by a boundary drawn on the world map that encloses an area delimited by naturally occurring geographical features (for example, river basins, coastal plains, valleys, and islands).

Within each world region, we looked for a set of NGAs that would allow us to cover as wide a range of forms of social organization as possible. Accordingly, we selected three NGAs that varied in the antiquity of centralized, stratified societies (giving us one early-complexity, one late-complexity, and one intermediate-complexity NGA per region).

Our unit of analysis is a polity, an independent political unit that ranges in scale from villages (local communities) through simple and complex chiefdoms to states and empires. To code social complexity data, for each NGA, our team chronologically listed all polities that were located in the NGA or encompassed it (*SI Appendix, SI Methods* has a discussion of how we deal with cases where identifying a single polity is not appropriate). For each NGA, we start at a period just before the Industrial Revolution (typically 1800 or 1900 CE depending on the location) and go back in time to the Neolithic (subject to the limitation of data). We chose a temporal sampling rate of 100 y, meaning that we only included polities that spanned a century mark (100, 200 CE, etc.) and omitted any polities of short duration that only inhabited an NGA between these points. Data collection was accomplished by a team of research assistants guided by archaeologists and historians who are experts in the sampled regions and time periods. These experts also checked all data collected by research assistants. *SI Appendix, SI Methods* contains details about coding procedures, including how we decided on the variables to include in the Seshat codebook and how we explicitly engaged with such issues as missing data, uncertainty, and disagreement between experts. We have created a website (seshatdatabank.info/) that illustrates the databank. This online version currently displays information on the social complexity variables in the NGAs and polities analyzed in this study (see also *SI Appendix, SI Methods*). The website shows how entries in the databank are supplemented by explanations of coding decisions and references. The goal of the databank is to make as explicit as possible the evidentiary basis of inferences about the past and to share that information as widely as possible.

Multiple Imputation: Dealing with Missing Data, Uncertainty, and Expert Disagreement. Because of the fragmentary nature of the information that is available about past societies, it was not possible to reliably code all variables for all polities. There is, therefore, a nontrivial amount of data points for which we have been unable to assign even a broad range of possible values because of a lack of evidence (3,700 of the total of 21,000). The presence of such missing data is an important feature of our dataset, in that it accurately reflects our current understanding (or lack of it) about any particular feature in any particular past society. Missing data, however, present a challenge for the statistical analyses.

One way of dealing with incomplete datasets is to simply omit the rows in the data matrix that contain missing values. There are two problems with this approach. First, it can be very wasteful in that omitted rows may contain much useful information relating to the variables that we were able to code. Had we used this approach with our social complexity data, for example, we would have to throw away nearly one-half of the rows. Second, case deletion may lead to biased estimates, because there are often systematic differences between the complete and incomplete cases. In our case, in many NGAs, small-scale societies were present far back in time, and as a result, they are much harder to code. Additionally, some regions of the world have been subject to greater levels of research effort than others. Omitting many of the lesser known cases because of their larger proportion of missing values would give too much weight to later, better known societies from only some parts of the world. As an example, had we used the casewise deletion approach for our current dataset, we would end up with only a single observation for Australia–Oceania. Such unequal dropping of observations would very likely bias the results, since the analysis would be dominated by such regions as Europe and southwest Asia (each with ~40 complete rows in the data matrix).

To deal with missing values as well as incorporate uncertainty and expert disagreement into our analyses, we use a technique known as multiple imputation (63), which utilizes modern computing power to extract as much information from the data as possible. Imputation involves replacing missing entries with plausible values, and this allows us to retain all cases for the analysis. A simple form of imputation, “single imputation,” might replace any unknown cases for a binary “present/absent” variable with simply “absent” or to replace unknown cases of continuous variables with the mean for that variable. These approaches have similar drawbacks to case deletion, in that they tend to introduce a bias. Therefore, in this paper, we perform multiple imputation: analysis done on many datasets, each created with different imputed values that are sampled in probabilistic manner. This approach results in valid statistical inferences that properly reflect the uncertainty caused by missing values (64). Multiple imputation procedures can vary depending on the type of variable and the type of data coding issue faced.

Expert disagreement. In cases where experts disagree, each alternative coding has the same probability of being selected. Thus, if there are two conflicting codings presented by different experts and if we create 20 imputed sets, each alternative will be used roughly 10 times.

Uncertainty. Values that are coded with a confidence interval are sampled from a Gaussian distribution, with mean and variance that are estimated assuming that the interval covers 90% of the probability. For example, if a value of [1,000–2,000] was entered for the polity population variable, we would draw values from a normal distribution centered on 1,500 with an SD of 304. It is worth noting that this procedure means that, in 10% of cases, the value entered into the imputed set will be outside the data interval coded in Seshat. For categorical or binary variables, we sample coded values in proportion to the number of categories that are presented as plausible. For example, if our degree of knowledge does not allow us to tell whether a certain feature was present or absent at a particular time, then the imputed datasets will contain “present” for roughly one-half of the imputed sets and absent for roughly one-half of the sets.

Missing data. For missing data, we impute values as follows. Suppose that, for some polity, we have a missing value for variable A and coded values for variables B–H. We select a subset of cases from the full dataset, in which all values of A–H variables have values and build a regression model for A. Not all predictors B–H may be relevant to predicting A, and thus, the first step is selecting which of the predictors should enter the model (information on model selection is given below). After the optimal model is identified, we estimate its parameters. Then, we go back to the polity (where variable A is missing) and use the known values of predictor variables for this polity to calculate the expected value of A using the estimated regression coefficients. However, we do not simply substitute the missing value with the expected one (because as explained above, this is known to result in biased estimates). Instead, we sample from the posterior distribution characterizing the prediction of the regression model (in practice, we randomly sample the regression residual and add it to the expected value). We applied the same

approach to each missing value in the dataset, yielding an imputed dataset without gaps.

The overall imputation procedure was repeated 20 times, yielding 20 imputed sets that were used in the analyses below. The 20 imputed datasets are available online as [Dataset S1](#).

Statistical Analysis.

PCA. PCA was used to investigate the internal correlation structure characterizing the nine measures of social complexity. PCA was run on each imputed dataset to estimate the proportion of variance explained by each PC (PC1–PC9), component loadings (correlations between the original variables and the PCs), and the values of PCs for each polity. Because we have 20 sets of all of these results, we also report the confidence intervals associated with these estimates. Values for PC1 derived from the 20 imputed datasets are available online as [Dataset S2](#).

Cross-validation. For the multiple imputation to be a worthwhile procedure, we need to ascertain that the stochastic regression approach for predicting missing values actually yields better estimates than, for example, simply using the mean of the variable. To do this, we used a statistical technique known as *k*-fold cross-validation (28). In addition to this methodological issue, this cross-validation procedure allows us to address another substantive question, namely the extent to which the relationships between variables are consistent across different parts of the world. This is done by quantifying how well we can predict the value of a particular feature of a particular society based on known information about the values of other features in that society and the observed relationships between the known and the unknown variables in other societies.

Cross-validation estimates the true predictability characterizing a statistical model by splitting data into two sets. The parameters of the statistical model are estimated on the fitting set. Next, this fitted model is used to predict the data in the testing set. Because the prediction is evaluated on the “out of sample” data (data that were not used for fitting the model), the results of the prediction exercise give us a much better idea of how generalizable the model is compared with, for example, such regression statistics as the coefficient of determination, R^2 .

The accuracy of prediction is often quantified with the coefficient of prediction (65):

$$\rho^2 = 1 - \frac{\sum_{i=1}^n (Y_i^* - Y_i)^2}{\sum_{i=1}^n (\bar{Y} - Y_i)^2},$$

where Y_i indicates the observations from the testing set (the omitted values), Y_i^* is the predicted value, \bar{Y} is the mean of Y_i , and n is the number of values to be predicted. The coefficient of prediction ρ^2 equals one if all data are perfectly predicted and zero if the regression model predicts as well as the data average (in other words, if the model is simply $Y_i^* = \bar{Y}$). Unlike the regression R^2 , which can vary between zero and one, prediction ρ^2 can be negative—when the regression model predicts data worse than the data mean. Prediction ρ^2 becomes negative when the sum of squares of deviations between predicted and observed is greater than the sum of squares of deviations from the mean.

In *k*-fold cross-validation, rather than having simply a single fitting set and one testing set, we divide the data into *k* sets. We selected those cases that had complete coding for all variables and divided our dataset into 10 sets for each of our 10 world regions. Next, we set aside one region (for example, Africa) and used the other nine regions to fit a regression model for the variable of interest. Let us say that Y is polity population, and we are interested in how well it can be predicted from knowing the population of the capital, hierarchy levels, writing, etc. We fit a regression model to the data from the other nine regions. We then predict the values of Y (polity population in this case) for Africa using the known values for other variables in African polities and the regression coefficients. Next, we omit another region (for example, Europe) and repeat the exercise. At the end, we have predicted all data points by the out of sample method, while fitting the model on 9/10th of data at any given step.

One important aspect of this procedure is to guard against overfitting (i.e., including too many predictor variables in the model), which is known to yield much worse predictability than a model that uses the “right” number of predictors (66). We have experimented with several methods of model selection that prevent overfitting. We found that a frequentist approach in which predictor variables are selected based on their *P* values (using the 0.05 threshold) does as well as the more commonly used model selection approach using the Akaike Information Criterion (AIC) (66). In fact, AIC tended

to slightly overfit compared with the frequentist approach. As the frequentist approach has an additional advantage of consuming less computer time, we used this approach for all cross-validation analyses reported below.

Multiple imputation, cross-validation, and PCA were all conducted using scripts written in the R statistical programming language (67).

ACKNOWLEDGMENTS. We thank Paula and Jerry Sabloff, Santiago Giraldo, and Carol Lansing who contributed to the development of Seshat. We also acknowledge Prof. Garrett Fagan, who passed away on March 11, 2017. He was a valued contributor to the Seshat Databank project, helping at an early stage in developing a coding scheme for social complexity variables and overseeing the coding of Roman polities. This work was supported by a John

Templeton Foundation Grant (to the Evolution Institute) entitled "Axial-Age Religions and the Z-Curve of Human Egalitarianism," a Tricoastal Foundation Grant (to the Evolution Institute) entitled "The Deep Roots of the Modern World: The Cultural Evolution of Economic Growth and Political Stability," Economic and Social Research Council Large Grant REF RES-060-25-0085 entitled "Ritual, Community, and Conflict," an Advanced Grant from the European Research Council under the European Union's Horizon 2020 Research and Innovation Programme Grant 694986, and Grant 644055 from the European Union's Horizon 2020 Research and Innovation Programme (ALIGNED; www.aligned-project.eu). T.E.C. is supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement 716212).

- Kohler TA, Crabtree SA, Bocinsky RK, Hooper PL (2015) Sociopolitical evolution in midrange societies: The pre-Hispanic Pueblo case. Available at <https://www.santafe.edu/research/results/working-papers/sociopolitical-evolution-in-midrange-societies-the>. Accessed May 24, 2017.
- Carneiro RL (2003) *Evolutionism in Cultural Anthropology* (Westview, Boulder, CO).
- Currie TE, Mace R (2011) Mode and tempo in the evolution of socio-political organization: Reconciling "Darwinian" and "Spencerian" evolutionary approaches in anthropology. *Philos Trans R Soc Lond B Biol Sci* 366:1108–1117.
- Spencer CS (1990) On the tempo and mode of state formation—Neoevolutionism reconsidered. *J Anthropol Archaeol* 9:1–30.
- White LA (1949) *The Science of Culture: A Study of Man and Civilization* (Farrar, Straus and Company, New York).
- Morris I (2013) *The Measure of Civilization: How Social Development Decides the Fate of Nations* (Princeton Univ Press, Princeton).
- Chick G (1997) Cultural complexity: The concept and its measurement. *Cross Cult Res* 31:275–307.
- Ortman SG, Cabaniss AHF, Sturm JO, Bettencourt LMA (2015) Settlement scaling and increasing returns in an ancient society. *Sci Adv* 1:e1400066.
- Currie TE, Mace R (2009) Political complexity predicts the spread of ethnolinguistic groups. *Proc Natl Acad Sci USA* 106:7339–7344.
- Spencer CS, Redmond EM (2001) Multilevel selection and political evolution in the Valley of Oaxaca, 500–100 BC. *J Anthropol Archaeol* 20:195–229.
- Turchin P, Currie TE, Turner EAL, Gavrilets S (2013) War, space, and the evolution of Old World complex societies. *Proc Natl Acad Sci USA* 110:16384–16389.
- Blanton RE, Feinman GM, Kowalewski SA, Peregrine PN (1996) A dual-processual theory for the evolution of Mesoamerican civilization. *Curr Anthropol* 37:1–14.
- Keech McIntosh S (1999) *Beyond Chiefdoms: Pathways to Complexity in Africa*, ed Keech McIntosh S (Cambridge Univ Press, Cambridge, UK), pp 1–30.
- Feinman G (2013) *Cooperation and Collective Action: Archaeological Perspectives*, ed Carballo DM (Univ Press of Colorado, Boulder, CO), pp 35–56.
- Yoffee N (2005) *Myths of the Archaic State: Evolution of the Earliest Cities, States, and Civilizations* (Cambridge Univ Press, Cambridge, UK).
- Hunt L (1989) *The New Cultural History* (Univ of California, Berkeley, CA).
- Sheehan JJ (2005) President's Column: How do we learn from history? *Perspect Hist* 43:1–3.
- Feinman GM (2008) Variability in states: Comparative frameworks. *Soc Evol Hist* 7:54–66.
- Turchin P, Gavrilets S (2009) Evolution of complex hierarchical societies. *Soc Evol Hist* 8:167–198.
- Spencer CS (2010) Territorial expansion and primary state formation. *Proc Natl Acad Sci USA* 107:7119–7126.
- Drennan RD, Peterson CE (2006) Patterned variation in prehistoric chiefdoms. *Proc Natl Acad Sci USA* 103:3960–3967.
- Feinman GM, Neitzel J (1984) *Advances in Archaeological Method and Theory*, ed Schiffer MB (Academic, Orlando, FL), Vol 7, pp 39–102.
- Currie TE (2014) Developing scales of development. *Cliodynamics* 5:65–74.
- Peregrine PN, Ember CR, Ember M (2004) Universal patterns in cultural evolution: An empirical analysis using Guttman scaling. *Am Anthropol* 106:145–149.
- Peregrine P, Ember CR, Ember M (2007) Modeling state origins using cross-cultural data. *Cross-Cultural Res* 41:75–86.
- Turchin P, et al. (2015) Seshat: The Global History Databank. *Cliodynamics* 6:77–107.
- Eff EA, Does MR (2004) Galton still have a problem? Autocorrelation in the standard cross-cultural sample. *World Cult* 15:153–170.
- Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence* (Morgan Kaufmann Publishers Inc., San Francisco), Vol 2, pp 1137–1145.
- Currie TE, Greenhill SJ, Gray RD, Hasegawa T, Mace R (2010) Rise and fall of political complexity in island South-East Asia and the Pacific. *Nature* 467:801–804.
- Diamond J (1997) *Guns, Germs and Steel* (Vintage, London).
- Gavrilets S (2015) Collective action problem in heterogeneous groups. *Philos Trans R Soc Lond B Biol Sci* 370:20150016.
- Mattison SM, Smith EA, Shenk MK, Cochrane EE (2016) The evolution of inequality. *Evol Anthropol* 25:184–199.
- Richerson PJ, Boyd R (2001) *The Origin of Human Social Institutions*, ed Runciman WG (Oxford Univ Press, Oxford), pp 197–234.
- Henrich J (2015) *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter* (Princeton Univ Press, Princeton).
- Marcus J (1998) The peaks and valleys of ancient states: An extension of the dynamic model. *Archaic States* (School of American Research Press, Santa Fe, NM), pp 59–94.
- Denham WW, McDaniel CK, Atkins JR (1979) Aranda and Alyawara kinship: A quantitative argument for a double helix model. *Am Ethnologist* 6:1–24.
- Cook M (2003) *A Brief History of the Human Race* (Granta Books, London).
- Mathew S, Boyd R (2011) Punishment sustains large-scale cooperation in prestate warfare. *Proc Natl Acad Sci USA* 108:11375–11380.
- Dunbar RIM (1995) *The Trouble with Science* (Faber & Faber, London).
- Collins R (1994) *Sociological Forum* (Springer, Berlin), Vol 9, pp 155–177.
- Wright HT (1977) Recent research on the origin of the state. *Annu Rev Anthropol* 6:379–397.
- Crumley CL (1995) *Heterarchy and the Analysis of Complex Societies*, eds Ehrenreich RM, Crumley CL, Levy JE (Archaeological Papers of the American Anthropological Association, Arlington, VA), Vol 6, pp 1–5.
- Blanton R, Fargher L (2008) *Collective Action in the Formation of Pre-Modern States* (Springer, Berlin).
- Carballo DM, Roscoe P, Feinman GM (2014) Cooperation and collective action in the cultural evolution of complex societies. *J Archaeol Method Theory* 21:98–133.
- Bondarenko DM, Grinin LE, Korotayev AV (2002) Alternative pathways of social evolution. *Soc Evol Hist* 1:54.
- Bellah RN (2011) *Religion in Human Evolution: From the Paleolithic to the Axial Age* (Harvard Univ Press, Cambridge, MA).
- Turchin P (2016) *Ultrasociety: How 10,000 Years of War Made Humans the Greatest Cooperators on Earth* (Beresta Books, Chaplin, CT).
- Wright HT (1986) *American Archeology Past and Future*, ed Meltzer DJ (Smithsonian Institution, Washington, DC), pp 323–365.
- Drennan RD (1991) Pre-hispanic chiefdom trajectories in Mesoamerica, Central America, and northern South America. *Chiefdoms: Power, Economy, and Ideology*, ed Earle TK (Cambridge Univ Press, Cambridge, UK), p 263.
- Kirch PV (1984) *The Evolution of the Polynesian Chiefdoms*, New Studies in Archaeology (Cambridge Univ Press, Cambridge, UK).
- Claessen HJM, Skalnik P (1978) *The Early State* (Moulton Publishers, The Hague, The Netherlands).
- Trigger BG (2003) *Understanding Early Civilizations: A Comparative Study* (Cambridge Univ Press, Cambridge, UK).
- Johnson AW, Earle T (2000) *The Evolution of Human Societies* (Stanford Univ Press, Stanford, CA).
- Flannery K, Marcus J (2012) *The Creation of Inequality: How Our Prehistoric Ancestors Set the Stage for Monarchy, Slavery, and Empire* (Harvard Univ Press, Cambridge, MA).
- Spencer CS (1997) Evolutionary approaches in archaeology. *J Archaeol Res* 5:209–264.
- Kirch PV, Green RC (1997) History, phylogeny, and evolution in Polynesia. *Curr Anthropol* 33:161–186.
- Spencer CS (2009) *Macroevolution in Human Prehistory*, eds Prentiss AM, Kuijt I, Chatters JC (Springer, New York), pp 133–156.
- McShea DW (1994) Mechanisms of large-scale evolutionary trends. *Evolution* 48:1747–1763.
- McShea DW (2001) The minor transitions in hierarchical evolution and the question of a directional bias. *J Evol Biol* 14:502–518.
- Carneiro RL (1970) A theory of the origin of the state: Traditional theories of state origins are considered and rejected in favor of a new ecological hypothesis. *Science* 169:733–738.
- Currie TE, et al. (2015) Agricultural productivity in past societies: Toward an empirically informed model for testing cultural evolutionary hypotheses. *Cliodynamics* 6:24–56.
- Whitehouse H, François P, Turchin P (2015) The role of ritual in the evolution of social complexity: Five predictions and a drum roll. *Cliodynamics* 6:199–216.
- Rubin DB (1987) *Multiple Imputation for Nonresponse in Surveys* (Wiley, New York).
- Yuan Y (2000) *Multiple Imputation for Missing Data: Concepts and New Developments* (SAS Institute, Rockville, MD), p 267.
- Turchin P (2003) *Complex Population Dynamics: A Theoretical/Empirical Synthesis* (Princeton Univ Press, Princeton).
- Burnham KP, Anderson DR (2002) *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach* (Springer, New York).
- R Core Team (2015) R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna).