

Explaining marriage patterns in a globally representative sample through socio-ecology and population history: A Bayesian phylogenetic analysis using a new supertree

Riana Minocher^{a,b}, Pavel Duda^c, Adrian V. Jaeggi^{a,1,*}

^a Department of Anthropology, Emory University, 1557 Dickey Drive, Atlanta, GA 30322, USA

^b Department of Human Behavior, Ecology, and Culture, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, Leipzig 04103, Germany

^c Department of Zoology, Faculty of Science, University of South Bohemia, České Budějovice, Czechia

ARTICLE INFO

Keywords:

Cultural evolution
Marriage
Phylogeny
Polygyny
Sexual selection
Standard Cross-Cultural Sample

ABSTRACT

Comparative analyses have sought to explain variation in human marriage patterns, often using predictions derived from sexual selection theory. However, most previous studies have not controlled for non-independence of populations due to shared ancestry. Here we leverage a phylogenetic supertree of human populations that includes all 186 populations in the Standard Cross-Cultural Sample (SCCS), a globally representative and widely-used sample of human populations. This represents the most comprehensive human phylogeny to date, and allows us not only to control for non-independence, but also to quantify the role of population history in explaining behavioral variation, in addition to current socio-ecological conditions. We use multiple imputation to overcome missing data problems and build a comprehensive Bayesian phylogenetic model of marriage patterns with two correlated response variables and eleven minimally collinear predictors capturing various socio-ecological conditions. We show that ignoring phylogeny could lead to both false positives and false negatives, and that the phylogeny explained about twice as much variance as all the predictors combined. Pathogen stress and assault frequency emerged as the predictors most strongly associated with polygyny, which had been considered evidence for female choice of good genes and male intra-sexual competition or male coercion, respectively. Mixed support was found for a polygyny threshold based on variance in male wealth, which is discussed in light of recent theory. Barring caveats, these findings refine our understanding of the evolution of human marriage systems, and highlight the value of combining population history and current socio-ecology in explaining human behavioral variation. Future studies using the SCCS should do so using the present supertree.

1. Introduction

1.1. Explaining variation in human marriage patterns

Marriage patterns vary widely within and across human societies, with most societies allowing polygyny yet most marriages being monogamous (Marlowe, 2003; White, 1988), calling for a comprehensive explanatory framework. Sexual selection theory accounts for the distribution of mating patterns across species (Clutton-Brock & Vincent, 1991; Emlen & Oring, 1977; Kokko & Jennions, 2008; Trivers, 1972), and may thus provide insights to human marriages. For instance, polygynous mating is prevalent when variance in male quality is high,

such that a few males can either directly control access to multiple females and exclude other males, or offer better genes or more resources relevant for female fitness (Emlen & Oring, 1977). In the latter case, females may choose to mate polygynously rather than monogamously if this offers greater expected fitness benefits, which is known as the polygyny threshold model (Borgerhoff Mulder, 1988, 1990). Conversely, monogamous mating may prevail in male-biased or widely dispersed populations as a form of mate-guarding, or when there are high returns to male parental investment (Kokko & Jennions, 2008; Lukas & Clutton-Brock, 2013; Marlowe, 2000; Schacht & Borgerhoff Mulder, 2015). More recently, this theory has been expanded explicitly for humans to include mutual mate choice and distinguish between

* Corresponding author.

E-mail address: adrian.jaeggi@iem.uzh.ch (A.V. Jaeggi).

¹ Current affiliation: Institute of Evolutionary Medicine, University of Zurich, Winterthurerstr. 190, 8057 Zurich, Switzerland.

rival and non-rival forms of wealth,² in order to explain the paradoxical decline in polygyny in the most unequal, often agricultural societies (Fortunato, 2011; Fortunato & Archetti, 2010; Oh, Ross, Borgerhoff Mulder, & Bowles, 2016; Ross et al., 2018).

Evolutionary anthropologists have applied sexual selection theory to explain variation in human marriage patterns (see Table 1 for overview of theoretical models and empirical tests). In support of the polygyny threshold model based on male wealth (*model 1a*), Kipsigis women chose husbands that offered better resources at the time, as measured by a (co-)wife's expected share of the husband's land (Borgerhoff Mulder, 1990). Polygyny among BaYaka was not clearly explained by good genes (Chaudhary et al., 2015; *model 1b*), and polygyny in Australian aborigines was better explained by male coercion (*model 2*) than female choice (Chisholm & Burbank, 1991). Others found evidence for market effects (*model 3*) as male mating effort (Schacht & Borgerhoff Mulder, 2015) and polygyny (Pollet & Nettle, 2009) were higher in places with female-biased sex ratios. Similarly, comparative studies using samples such as the Ethnographic Atlas or the Standard Cross-Cultural Sample (SCCS, see Fig. 1) have found evidence for a polygyny threshold in that polygyny was higher in populations with greater male control over resources (Cowlshaw & Mace, 1996; Hartung, 1982; Sellen & Hruschka, 2004; *model 1a*), as well as higher pathogen stress and lower male contribution to subsistence (Low, 1988, 1990; Marlowe, 2000; *model 1b*). Support for harem-defense polygyny (*model 2*) was provided by measures of male coercion (such as arranged female marriages) or male-male competition (such as the frequency of warfare or assault; Ember, 1974; Ember, Ember, & Low, 2007; Marlowe, 2003; White & Burton, 1988). Comparative tests have provided mixed associations of polygyny with processes influencing the adult sex ratio (Ember, 1974; Quinlan & Quinlan, 2007; *model 3*). Some studies have also suggested that marriage patterns are a product of societal organization and complexity (Osmond, 1965; Sanderson, 2001), with theorists providing different explanations for the decrease in polygyny in the most complex societies (*models 4 and 5*). However, these comparative studies are limited by several methodological difficulties including non-independence and missing data, and they fail to account for the influence of population history on current behavior, as postulated e.g. by cultural evolution theory (Richerson & Boyd, 2004).

1.2. Problems and opportunities with comparative analyses

Data points in comparative datasets, whether species or populations, cannot be assumed to be statistically independent because related groups may share similar traits due to common ancestry (Felsenstein, 1985; Mace & Pagel, 1994). This can lead to elevated Type I and Type II error rates (Harvey & Rambaut, 1998; Lindenfors, Revell, & Nunn, 2010), even in datasets designed to maximize cultural independence such as the SCCS (Dow, 1993; Dow & Eff, 2008; Murdock & White, 1969). This can be avoided by controlling for similarity by descent using a phylogenetic tree, which reflects the expected covariance among related groups (Felsenstein, 1985; Mace & Pagel, 1994; Nunn, 2011). Thus, a number of recent cross-cultural analyses have used phylogenies based on linguistic or genetic data (e.g. Fortunato, Holden, & Mace, 2006; Jordan, Gray, Greenhill, & Mace, 2009; Mace & Sear, 2005; Sheehan, Watts, Gray, & Atkinson, 2018; Walker, Hill, Flinn, & Ellsworth, 2011). Relevant for the present topic, Fortunato (2011) reconstructed the ancestral marriage pattern among Indo-Europeans to be monogamy, and Cowlshaw and Mace (1996) re-analyzed and confirmed Hartung's (1982) associations of polygyny with bridewealth and

male-biased wealth inheritance. While representing a huge advance over previous comparative approaches in anthropology, most studies were limited to available linguistic phylogenies, and thus geographically restricted samples of populations that shared a last common ancestor quite recently (e.g. Indo-Europeans, Bantu, or Austronesians), or available genetic phylogenies, and thus small, potentially skewed samples. A few studies on larger samples combined several linguistic phylogenies (Cowlshaw & Mace, 1996; Moravec et al., 2018), but without formally integrating them into a single tree.³ In sum, there is a rich literature on phylogenetic analyses of regional samples, but by spanning a larger breadth of human environments and cultures, global samples arguably provide more power to detect independent co-evolutionary events, here convergent evolution of marriage patterns in response to similar socio-ecological conditions, which is what phylogenetic comparative methods were designed to do (Felsenstein, 1985; Mace & Pagel, 1994; Nunn, 2011).

To apply the phylogenetic approach to a globally representative cross-cultural sample, we here leverage a recently expanded supertree of human populations (Duda & Zrzavý, 2016; Duda & Zrzavý, *in press*) that combines genetic and linguistic data from 388 individual phylogenies into a single tree. Such a supertree has the advantage of being topologically robust due to the incorporation of many phylogenetically informative characters, and to allow branch lengths and divergence times to be estimated even among distantly related groups. While the robustness of the supertree topology and inferences for human population history are discussed elsewhere (Duda & Zrzavý, *in press*), we trimmed this tree to the 186 SCCS populations and dated it for the present study (Fig. 2; see 2.1., Supplementary Material for details on tree inference and time calibration). Thus, the present SCCS supertree is the first dated phylogeny of a globally representative sample of human populations, and with a root estimated at 140'000 years before present and populations from all over the world reflects much of the full depth and breadth of human history (see Figs. 1 & 2). In the future, the supertree could be expanded to even broader samples such as the Ethnographic Atlas (Murdock, 1967).

Note that several studies found evolutionary patterns, e.g. rates of evolution or associations between variables to differ on different language trees (Cowlshaw & Mace, 1996; Moravec et al., 2018), which arguably speaks against using a single tree to represent all human populations. However, varying evolutionary patterns can in principle be accounted for by a single phylogeny, e.g. through random slope models, which has the added benefit that parameter estimates can inform each other (i.e. “pooling”, McElreath, 2016), leading to more robust inference, and that evolutionary patterns can vary continuously across the tree, rather than being fixed within certain clades. Unfortunately, random slope versions of our models were not feasible to implement here. In addition, future models could distinguish between predictors of the origin and maintenance of marriage patterns (Hansen, 2014; Ross, Strimling, Erickson, Lindenfors, & Borgerhoff Mulder, 2016). In the absence of such extensions, we note that our analysis (i) might misestimate associations between marriage patterns and socio-ecological predictors if these differ across clades, and (ii) only detects predictors of the origins but not maintenance of marriage patterns.

In sum, the SCCS supertree allows phylogenetic analyses to be applied for the first time to the most widely used cross-cultural sample, thus increasing our power to test adaptive hypotheses. The present study also serves to critically appraise the previous SCCS studies mentioned above, which did not control for phylogenetic non-independence and could thus be vulnerable to statistical artifacts.

² Rival forms of wealth diminish in value when shared among multiple people, e.g. land or livestock have to be divided when passed on to several offspring. Non-rival forms of wealth do not diminish in value, e.g. a male's genes, knowledge, or protection can equally benefit several offspring

³ This is because linguistic characters such as cognates, while ideal for quantifying similarity among recently diverged populations, cannot easily be used to relate language families to one another because they evolve too quickly (but see Pagel, Atkinson, S. Calude, & Meade, 2013); in our supertree genetic data provide the necessary deep history

Table 1

Theoretical models explaining the distribution of marriage patterns in humans, their logic and predictions, and major empirical tests.

Theoretical model ^a	Main logic/predictions	Empirical tests, single population	Empirical tests, comparative ^b
1: <i>Polygyny threshold</i> 1a: Based on male wealth, i.e. <i>resource-defense polygyny</i>	Polygyny greater with greater variance in male resource control	- Kipsigis, Kenya (Borgerhoff Mulder et al., 1990): women chose husbands that provided greatest expected share of resources (livestock) at the time, and male land ownership predicts number of wives	- WNAI (Sellen & Hruschka, 2004): variance in male reproductive success associated with male resource control - Various (Betzig, 2012): Variance in male reproductive success among agriculturalists > > pastoralists/ horticulturalists > foragers - EA (Cowlshaw & Mace, 1996, Hartung, 1982): polygyny associated with brideprice and male-biased inheritance
1b: Based on male genetic quality	Polygyny, i.e. female preference for good genes, greater if need for genetic quality high, or need for paternal investment low	- BaYaka, Congo (Chaudhary et al., 2015): polygynous males not better phenotypic quality, but higher social capital	- SCCS (Low, 1990): degree of polygyny increases with pathogen stress - SCCS (Marlowe, 2000): Polygyny higher when paternal care (direct proximity to infants and contribution to subsistence) is low
2: <i>Harem-defense polygyny</i> , via male-male competition and male coercion	Polygyny associates with male-male competition (e.g. violence), and arrangements facilitating male alliances (e.g. patriline, patrilocality/ virilocality/ avunculocality)	- Arnhem Land, Australia (Chisholm & Burbank, 1991): Polygyny better explained by male coercion than female choice	- SCCS foragers (Marlowe, 2003): Polygyny associated with arranged female marriages and assault frequency - SCCS (Carter & Kushnick, 2018; Marlowe, 2000): Polygyny associated with male aggressiveness - HRAF (Betzig, 1982): Number of wives for leaders correlates with despotism - SCCS (White & Burton, 1988): Polygyny associated with fraternal interest groups
3: <i>Market forces</i> , polygyny and paternal investment depend on sex ratio	Polygyny greater and paternal investment lower when sex-ratio female-biased as males are in high demand and costs of finding another female are low	- Uganda (Pollet & Nettle, 2009): frequency of polygyny is greater in female-biased districts, independent of wealth - Makushi communities, Guyana (Schacht & Borgerhoff Mulder, 2015): male mating effort increases in male-biased villages	- SCCS (Quinlan & Quinlan, 2007): Polygyny associated with pair bond stability, because polygyny creates a shortage of female partners - EA (Ember, 1974; Ember, 1984): Polygyny associated with female-biased sex ratios, and male mortality during warfare (which biases sex-ratio to females) - SCCS (White & Burton, 1988): Polygyny associated with factors affecting adult sex ratio, e.g. warfare for plunder and marriage of captive women, subsistence type and homogeneity of landscape - SCCS (Ember et al., 2007): male mortality during warfare predicts higher polygyny in nonstate societies, where expected to affect sex-ratio more than in state societies
4: <i>Paternal investment of rival wealth</i> (Fortunato & Archetti, 2010; Oh et al., 2016)	Lower polygyny when high paternal investment required but resources are rival (e.g. land)		- Individual-level data from 29 societies (Ross et al., 2018): Low levels of and decreasing marginal returns to polygyny in stratified agriculturalists despite high wealth inequality - Indo-europeans (Fortunato, 2011): Origins of monogamy better explained by agriculture than societal complexity
5: <i>Cultural group selection of monogamy</i> (Alexander, 1979; Henrich et al., 2012)	Polygyny decreases with increasing societal complexity, normative monogamy associated with lowered levels of male-male competition (e.g. crime)	- Mormons (cited in Henrich et al., 2012): Decline in strength of intra-sexual selection after imposed monogamy - India, China (cited in Henrich et al., 2012): Increasing crime with increasingly male-biased sex ratio	- EA (Osmond, 1965): marriage type influenced by complexity of social organization - SCCS (Sanderson, 2001): socially imposed monogamy more common in larger states - Various countries (Henrich et al., 2012): Higher infant and child mortality in polygynous countries - Various (Schacht et al., 2014): Mixed evidence for violent male-male competition when sex ratio is male-biased

^a References given only for models specifically designed to explain human marriage.^b WNAI = Western North American Indians, EA = Ethnographic Atlas, SCCS = Standard Cross-Cultural Sample.

Incorporating a phylogeny in comparative analyses also allows us to extend the explanatory scope beyond current socio-ecological conditions to include population history. Adaptive hypotheses typically assume that behavioral variation is patterned by current socio-ecological conditions; in other words, individuals adopt optimal strategies given costs and benefits in the current environment (Nettle, Gibson, Lawson, & Sear, 2013; Winterhalder & Smith, 2000). However, it is also well known that phylogenetic history may limit behavioral variation in a range of species (Shultz, Opie, & Atkinson, 2011; Thierry, 2008), either due to constraints on adaptation or because past adaptations are

preserved through stabilizing selection (Hansen, 2014). In humans, various processes could make current practices dependent on population history, including genetic inheritance, vertical and horizontal cultural transmission, or niche conservatism (Borgerhoff Mulder, George-Cramer, Eshleman, & Ortolani, 2001; Richerson & Boyd, 2004), allowing the use of phylogenies to reconstruct ancestral states in the first place (Fortunato, 2011; Jordan et al., 2009; Opie, Shultz, Atkinson, Currie, & Mace, 2014; Pagel, 1999). Here we use tools borrowed from quantitative genetics (de Villemereuil & Nakagawa, 2014; Hadfield & Nakagawa, 2010) to partition variance in behavior into components

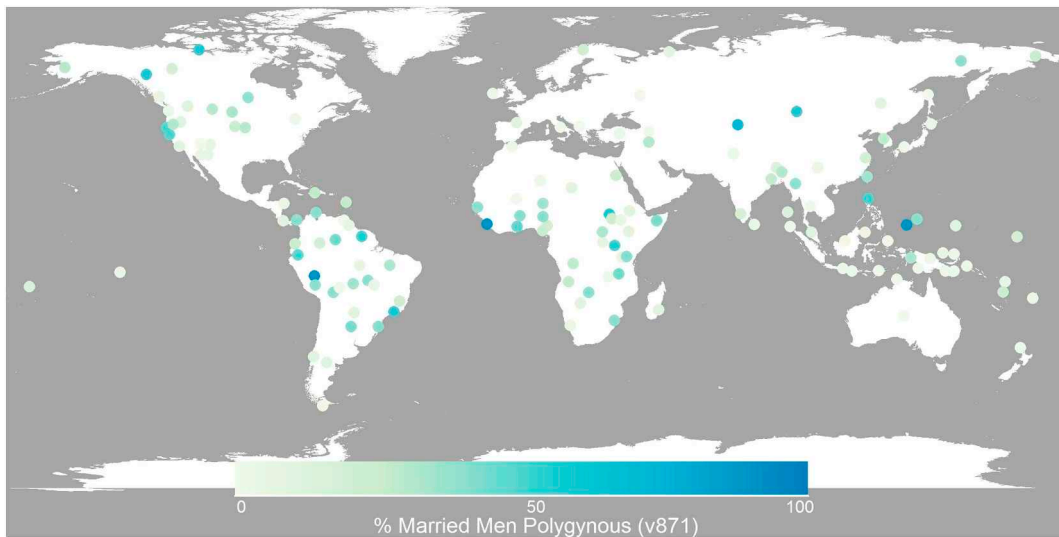


Fig. 1. World map showing the distribution of all SCCS populations with information on the percent of married men who are polygynous (v871).

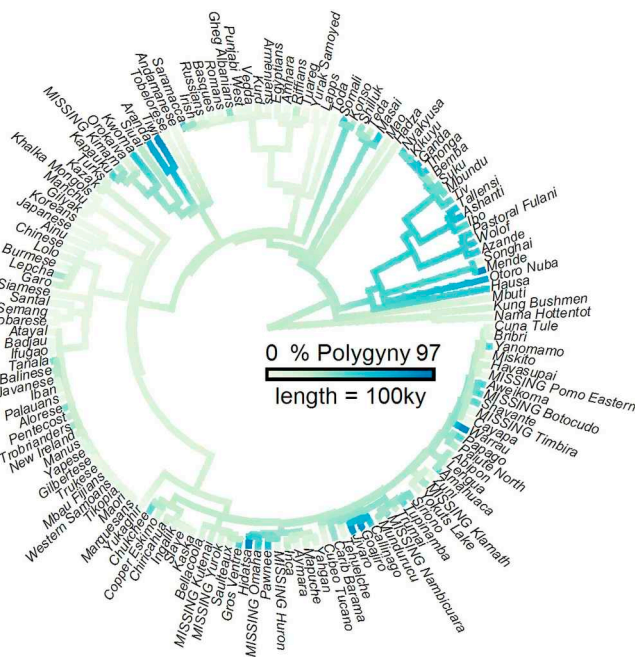


Fig. 2. The SCCS supertree showing all populations with information on the percent of married men who are married to more than one woman (v871). The influence of population history on current levels of polygyny is evident: entire clades resemble each other in color (e.g. East Asian and Pacific groups on the left have mostly low levels of polygyny, while many African populations on the lower right have high levels of polygyny). Internal branches are color-coded by inferred ancestral state, as implemented in the *phytools* package version 0.6–44. The last common ancestors of all extant human populations are estimated to have had moderate levels of polygyny.

due to measurable aspects of the current environment, i.e. socio-ecological predictors, and population history, i.e. the phylogeny (see Section 2.2.). The latter potentially captures any of the aforementioned processes, as well as unmeasured socio-ecological similarity between related groups (similar to heritability estimates in quantitative genetics arguably containing unmeasured environmental signals (Joseph, 2014)). This method represents a significant methodological advance over previous comparative studies that have contrasted current ecology and population history as predictors of human behavioral variation in smaller geographic samples (Borgerhoff Mulder et al., 2001;

Guglielmino, Viganotti, Hewlett, & Cavalli-Sforza, 1995; Hewlett, De Silvestri, & Guglielmino, 2002; Mathew & Perreault, 2015). We also caution, along with others (Towner, Grote, & Borgerhoff Mulder, 2015), against equating variance explained by population history (a pattern) with cultural transmission (one of several possible processes consistent with this pattern). While our analysis does not account for geographic distance in addition to phylogenetic distance, the two are highly correlated in the SCCS and any processes expected to lead to behavioral similarity based on geographic distance (such as horizontal cultural transmission, or unmeasured ecological similarity) are thus subsumed by the phylogeny. Indeed, models including both the supertree and geographic distance were difficult to estimate in a related study due to high collinearity between the two (Ringen & Jaeggi, 2018); in other words, it may not be possible to partial out the influence of shared history from that of shared geography in the SCCS (see Borgerhoff Mulder, Nunn, & Towner, 2006 for a general discussion of the difficulty of inferring vertical vs horizontal transmission of cultural traits, including for polygyny in East Africa). Thus, our analyses account for behavioral variation patterned by current conditions as well as population history.

Finally, modern statistical methods allow researchers to improve upon previous work by making better use of existing data, which is especially important in the invaluable but patchy ethnographic record. In particular, virtually all previous comparative studies of marriage patterns cited above used complete-case analyses, wherein any populations with missing data were excluded (Dow & Eff, 2009). This drastically reduces the amount of data, and makes it impossible to combine multiple predictors of interest effectively as very few populations will have information on all variables of interest. As a result, many cross-cultural analyses only consider a small number of variables (Ember, 1974; Ember et al., 2007; Low, 1990; Marlowe, 2000; Marlowe, 2003; Quinlan & Quinlan, 2007; White & Burton, 1988), and present several analyses on different subsets of the data, depending on which variables were included. In addition to impeding broader synthesis, this practice effectively assumes that data are missing at random, otherwise excluding missing cases would bias parameter estimates. A better way to handle data assumed to be missing at random is through imputation, wherein missing records are replaced by a probability distribution or multiple predicted values (Carter & Kushnick, 2018; Dow & Eff, 2009; McElreath, 2016; Zhou & Reiter, 2010). In a Bayesian context, the uncertainty inherent in such imputation is easily carried forward and reflected in the final posterior distributions of parameter estimates; thus, more missing data simply lead to more uncertainty. We also make use of multi-response models (Hadfield, 2010),

Table 2
Description of variables. See SOM Table S1 and *Selecting predictor variables* for more details on all variables.

Variable	N	Prediction ^a	Model ^b	Related variables (see SOM)	Original coders
Measures of marriage system					
Cultural rules constraining polygyny (v860)	184			Polygamy code (v861)	(White, 1988)
% of married men polygynous (v871)	145			% women married polygynously (v872)	(White, 1988)
Predictors					
Role of father, infancy (v53)	154	–	1b		(Barry III & Paxson, 1971)
Population density (v64)	184	None	Control		(Murdock & Wilson, 1972)
Stratification (v158)	186	+	1a, 4, 5		(Murdock & Provost, 1973)
Temperature (v186)	180	None	Control	Region (v200), climate (v857), Latitude (v179, v180)	(Murdock, 1967; White, Whiting, & Burton, 1986; Whiting, Sodergren, & Stigler, 1982)
Sex ratio (v714)	90	–	3	Sex ratio (v1689)	(Ember & Ember, 1992)
Arranged female marriages (v740)	151	+	2		(Broude & Greene, 1983)
Female contribution to agriculture (v890)	185	+	1b	Various subsistence measures (see SOM)	(White, from: Barry III & Schlegel, 1982; Murdock, 1967; Whyte, 1985)
Pathogen stress (v1260)	186	+	1b		(Low, 1988)
Internal warfare (v1649)	152	+	2, 3	High value on male aggression (v625), male mortality from warfare	(Ember & Ember, 1992; Whyte, 1978); (Ember et al., 2007)
Assault frequency (v1666)	113	+	2		(Ember & Ember, 1992)
Wealth inequalities (v2021)	57	+	1a		(Pryor, 2003)

^a Predicted direction of association with polygyny. See Methods, Table S1 and supplied R code for details on coding and transformations.

^b See Table 1, Control = Control variable, no clear model or prediction.

which allow us to include several correlated outcome variables and pool information across them. Thus, imputation and multi-response models make better use of existing data, and we can effectively synthesize previously published studies on marriage in the SCCS (Ember et al., 2007; Low, 1990; Marlowe, 2000; Marlowe, 2003; Quinlan & Quinlan, 2007; Sanderson, 2001; White & Burton, 1988) by including all relevant predictors in the same model, in addition to incorporating phylogeny.

In sum, we report a Bayesian phylogenetic analysis of human marriage patterns, using two correlated outcome variables and eleven predictors used in previous SCCS studies that capture various socio-ecological conditions thought to be associated with degree of polygyny (see Table 2, Supplementary Material *Selecting predictor variables*). This is the most comprehensive cross-cultural analysis of marriage patterns to date, and the first phylogenetic analysis of a globally representative sample of human societies. We compare our phylogenetic model to a non-phylogenetic one, and then provide more detailed interpretations of the phylogenetic model. We show that (i) ignoring phylogeny leads to an elevated risk of both Type I and Type II errors, (ii) phylogeny explains substantial amounts of variance in polygyny, and (iii) the strongest predictors of polygyny are largely consistent with sexual selection accounts of human reproductive strategies, barring common caveats.

2. Materials & methods

2.1. Supertree

The principle of a supertree is to combine many smaller, partially overlapping source trees to obtain a single, comprehensive tree that encompasses an entire taxonomic group of interest (Bininda-Emonds, 2014), such as the comprehensive trees of living mammals (Bininda-Emonds et al., 2007) or birds (Jetz, Thomas, Joy, Hartmann, & Mooers, 2012), or, in this case, human populations. Here we introduce and use a supertree of the 186 populations present in the SCCS (Fig. 2), with details on supertree construction and time calibration provided in the Supplementary Materials. The supertree is an extension of a previously published tree (Duda & Zrzavý, 2016, in press) and combines 388 source trees based on 251 studies, including 12,770 phylogenetically informative characters – both genetic and linguistic – on 1979 human populations. The 186 populations in the SCCS thus represent a subset of a much larger tree, which improves the resolution and topological

stability. The SCCS supertree was time-calibrated using known time-constraints for 57 nodes.

2.2. Analysis approach

2.2.1. Outcome variables

There are several measures of polygyny available in the SCCS (White, 1988). We here focused on two variables (Table 2): the *cultural rules constraining the frequency of polygyny* (v860), and the *percentage of married men with more than one wife* (v871). The latter is highly correlated with the percentage of women polygynously married (v872; $r = 0.97$, $P < .0001$), which is arguably a better measure of the skew in the mating system and the intensity of sexual selection (Low, 1990), but v871 has fewer missing data and is sometimes deemed more reliable (Marlowe, 2003). In addition, previous studies have used various ordinal or binary measures of polygyny (Ember et al., 2007; Low, 1990; Marlowe, 2000; Marlowe, 2003), but we did not include these because they were either (i) constructed from v860 and/or v871/v872 and thus redundant (e.g. marriage codes used by (Low, 1990; Marlowe, 2000)), or (ii) created difficulty with model convergence and/or computation time (e.g. non-sororal polygyny (Ember et al., 2007)). Even though some values in v871 are implausibly high and v860 is crude, the combination of two outcomes in the same multi-response model allows us to place greater confidence in results that are consistent for both.

2.2.2. Predictor variables

Dozens of different variables have been considered by previous studies as predictors of polygyny in the SCCS (Ember et al., 2007; Low, 1990; Marlowe, 2000; Marlowe, 2003; Quinlan & Quinlan, 2007; Sanderson, 2001; White & Burton, 1988). These capture various aspects of subsistence, violence, social stratification, or climate, as well as other theoretically relevant features of the socio-ecology that do not easily group with these themes. Most predictors are thought to influence the degree of polygyny by capturing aspects of male-male competition or coercion, variance in male quality and thereby female choice, or returns to male investment (Table 2). Notably absent was a measure of religious beliefs influencing polygyny, as the SCCS (v713) lumps religions arguably promoting monogamy (Christianity) and polygyny (Islam); however, it can also be argued that societies adopt and adhere to such normative beliefs only insofar as they make sense in their socio-ecology (Osmond, 1965). In order to avoid overfitting the model, we reduced the total number of predictors and the collinearity between them (see SI

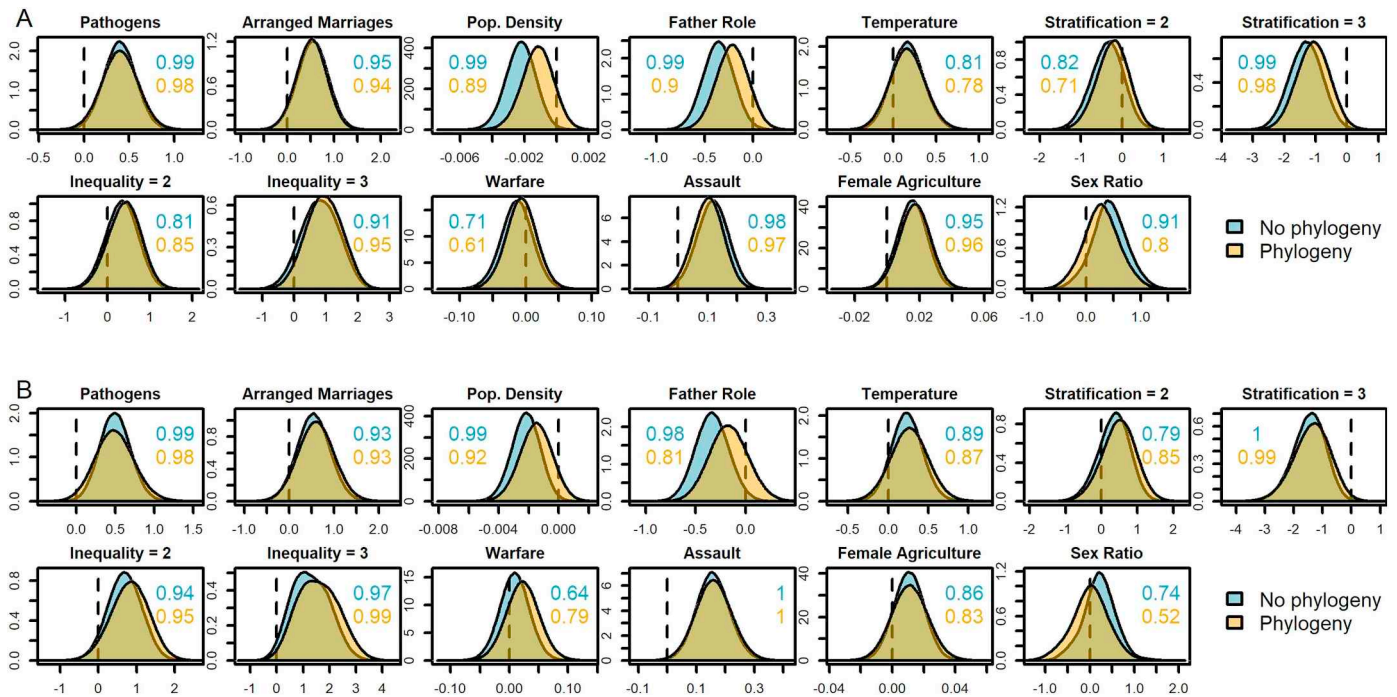


Fig. 3. Posterior probability distributions for all predictors of A) the percent of married men who are polygynous (v871) and B) cultural rules constraining polygyny (v860) from the non-phylogenetic (blue) and phylogenetic (yellow) models. The vertical dashed line at 0 indicates no association with the outcome. Numbers within each panel represent proportion of the distribution that is on the same side of 0 as the mean, i.e. the posterior probability that a predictor was associated with the outcome. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Selecting predictor variables for details), resulting in eleven predictors (Table 2, Table S1). It should be noted that one could endlessly add or remove candidate variables as the SCCS includes numerous, often partially redundant codes, each with its own pros and cons; we prefer to take a pragmatic approach here, by starting with previously used predictors and using objective variable-reduction methods – by making all data and R code available we invite readers to substitute their preferred variables into the analyses.

2.2.3. Data imputation and preparation

A dataset containing the two outcome measures and eleven predictors was compiled and two populations in which both outcomes were missing (Ajie, Gond) were excluded. From the remaining 184 cases, 100 complete datasets were imputed using *mice* version 2.30 (van Buuren & Groothuis-Oudshoorn, 2011). Such a large number of imputations allows MCMC samples from all individual models to be combined in order to describe the entire posterior distribution, accounting for the uncertainty in the imputation (Zhou & Reiter, 2010). Prior to fitting the model, predictor variables were transformed to facilitate model fit and interpretation (centered, or standardized) or reduce the number of factor levels (and therefore the number of parameters to be estimated). The factor levels were chosen to best capture meaningful variation in polygyny, based on inspecting boxplots of v871 against the variable of interest, with the least number of levels and thus model parameters, which should improve model predictions. Note that wealth inequalities (v2021) and social stratification (v158) were both converted to three-level factors, and could thus potentially capture the observed inverse-U shaped relationship between polygyny and social complexity (Oh et al., 2016; Ross et al., 2018). Details on these transformations can be found in the Supplementary Materials (Table S1) and the supplied R code.

2.2.4. Model fitting and summary

We used Bayesian phylogenetic multi-response models implemented in the *MCMCglmm* package (Hadfield, 2010) in R 3.2.3. (R Development

Core Team, 2015) to model both measures of polygyny simultaneously as correlated outcomes. This allows information about parameters to be pooled across the two outcomes, thus improving model accuracy (McElreath, 2016); such pooling is especially relevant for the phylogenetic variance component, which is often estimated with a high degree of uncertainty because each population or species only contributes a single observation.

The percent of married men who are polygynous (v871) was modelled with a Poisson distribution and a log link function, whereas the cultural rules constraining polygyny (v860) was modelled as an ordinal outcome with a probit link function. Ordinal models assume that the distance between levels is equal, which is a strong assumption – e.g. the distance between level 1 (monogamy prescribed) and level 2 (monogamy preferred) may be different from the distance between level 4 (polygyny for upper class) and level 5 (polygyny prevalent); ordered logistic models provide an alternative that relaxes this assumption (McElreath, 2016) but are not implemented in *MCMCglmm*. We fit the model both with and without the phylogeny. We used slightly regularizing priors to impose conservatism on the fixed effects (Gaussian distribution with mean = 0, SD = 5) and inverse gamma priors (scale parameter = 1, shape = 0.01 without phylogeny, shape = 1 with phylogeny) for the residual variance in the Poisson process and the phylogenetic variance and covariances; the residual variance for the ordinal outcome cannot be estimated and was fixed at 1 (Hadfield, 2016). For both the non-phylogenetic and phylogenetic model, we first checked convergence on a single dataset by checking trace plots of the Markov chains and effective sample sizes, and by calculating the Gelman-Rubin diagnostic using the *coda* package version 0.19–1 (Plummer, Best, Cowles, & Vines, 2006); convergence was considered sufficient when the upper confidence interval of the diagnostic was 1.01 or less and the effective sample size > 500 for all parameters. This required 500,000 iterations with a burnin of 1000 and a thinning interval of 10 for the non-phylogenetic model, and 100,000 iterations with a burnin of 10,000 and thinning interval of 10 for the phylogenetic model. These two models were then fit to each of the 100 imputed

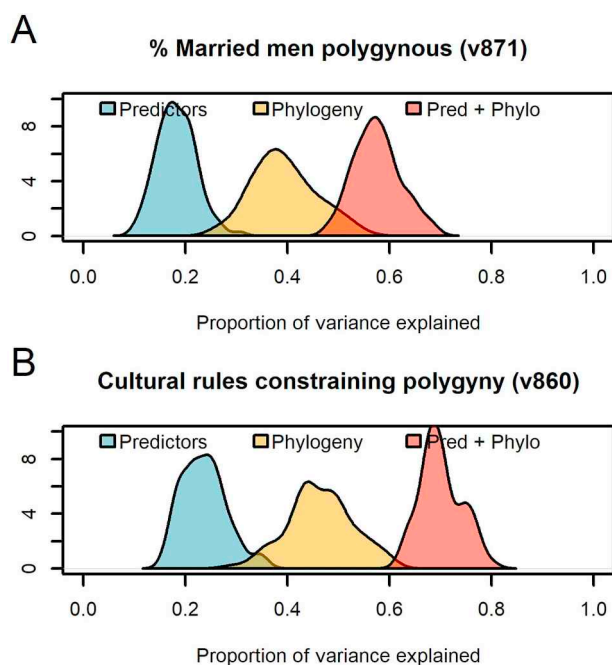


Fig. 4. The proportion of variance in A) v871 and B) v860 explained by predictors (blue), phylogeny (yellow), and both combined (red) for both measures of polygyny, across all 100 imputed datasets. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

datasets (which took several days for each model), and all MCMC samples from the 100 non-phylogenetic and 100 phylogenetic models respectively were combined to yield a single posterior distribution for each parameter.

To summarize the association between polygyny and predictors we present the posterior distributions of all regression coefficient estimates (Fig. 3), give the posterior probability (PP) of each predictor being associated with the outcome, i.e. the proportion of the distribution on the same side of 0 as the mean, and provide means and 95% credible intervals (Table S2). We also present predictive plots for the variables most strongly associated with polygyny (Figs. 5, 6, S1). To quantify the influence of phylogeny we calculate the intra-class correlation (Hadfield, 2016; Hadfield & Nakagawa, 2010), i.e. variance explained by phylogeny divided by total variance, which is identical to the commonly used Pagel's λ . In other words, the intra-class correlation captures the extent to which a population's marriage patterns can be predicted by knowledge of its phylogenetic relationships. Furthermore, we compare the proportion of total variance explained by phylogeny to that explained by the predictors, as well as both phylogeny and predictors combined (Fig. 4); these values are akin to the coefficient of determination R^2 (Nakagawa & Schielzeth, 2013) and provide a comparable measure of goodness of fit.

3. Results

Fig. 3 plots the posterior probability distribution for the association with polygyny for each predictor in both the phylogenetic (yellow) and non-phylogenetic (blue) models (see Table S2 for means and 95% credible intervals). The posterior probability (PP) is given in each figure panel. As expected, some predictors have reduced support in the phylogenetic model (e.g. population density, role of father in infancy, sex ratio), indicating an elevated risk of false positives when ignoring non-independence. However, others (e.g. wealth inequality, warfare) receive *more* support in the phylogenetic model, indicating an elevated risk of false negatives. In addition to visually checking for non-overlapping distributions, we can compare the PP's for phylogenetic and

non-phylogenetic models to ask which predictors would change from being considered significant to non-significant (or vice versa) under a traditional frequentist interpretation; the standard two-tailed significance threshold with $\alpha = 0.05$ would be $PP = 0.975$, since $> 1 - \alpha/2$ of the probability mass must be on the same side of 0 for the association to be considered significant. This threshold is crossed when including phylogeny for population density and role of father in infancy, which both become non-significant for both measures of polygyny, as well as for the association between level 3 inequality and cultural rules constraining polygyny, which becomes significant. Thus, this method identified four false positives and one false negative when failing to account for phylogeny.

In the phylogenetic model, the supertree captured a substantial amount of variance in both outcomes across all 100 imputed datasets. The intra-class correlation indicated a medium phylogenetic signal of 0.45 (95% Credible Interval = 0.18–0.73) for v871 and 0.56 (95% CI = 0.27–0.84) for v860. Nevertheless, the phylogeny on average captured more than twice as much variance as all the predictors combined, as shown in Fig. 4. Together, predictors and phylogeny explained about 60% and 70% of the variance in polygyny, for v871 and v860, respectively.⁴ The two outcomes also showed high covariance in their phylogenetic signal, with a mean correlation of 0.81 (95% CI = 0.55–0.98). Thus, our models provide strong evidence that population history influences human behavioral variation in a globally representative sample.

The socioecological predictors most strongly and consistently associated with polygyny in the phylogenetic model were pathogen stress (v1260, PP = 0.98 for both outcomes) and assault frequency (v1666, PP = 0.97 for v871 and 1.00 for v860), both of which increase polygyny. Thus, all else equal, only 1.63% of married men are expected to be polygynous in a population experiencing minimal pathogen stress, but 6.36% of married men in a population experiencing maximum pathogen stress; a four-fold increase (Fig. 5a). Similarly, at minimum levels of assault frequency 3.56% of married men are expected to be polygynous, but 8.05% at maximum level of assault frequency; a two-fold increase (Fig. 5b). The respective changes in the expected probability of cultural rules constraining polygyny are shown in Fig. 6 and Supplementary Material Fig. S1. Cultural rules prescribing monogamy are expected to decrease in probability from 0.29 to 0.05 and 0.11 to 0.02, i.e. almost six-fold, when going from minimal to maximal levels of pathogen stress and assault frequency, respectively. Conversely, the probability of a cultural rule for polygyny to be prevalent would increase from 0.01 to 0.13, and from 0.06 to 0.25, respectively. Other predictors were less clearly associated with polygyny, less consistent across the two outcome measures (e.g. female contribution to agriculture), or across different levels of the predictor (e.g. stratification, wealth inequality; but see 4.2.). However, we emphasize that Bayesian inference is probabilistic rather than based on arbitrary significance thresholds, and readers may draw their own conclusions about the support for various predictors based on the posterior probability distributions (Fig. 3).

4. Discussion

We used a phylogenetic supertree of human populations to appraise and synthesize previous studies of marriage patterns in the Standard Cross-Cultural Sample (SCCS). We first discuss the methodological aspects of this study and general implications for explaining human behavioral variation, and then highlight insights and caveats for our understanding of polygyny.

⁴ Note that for v860 the residual variance could not be estimated and this percentage is thus contingent on the value at which the residual variance was fixed.

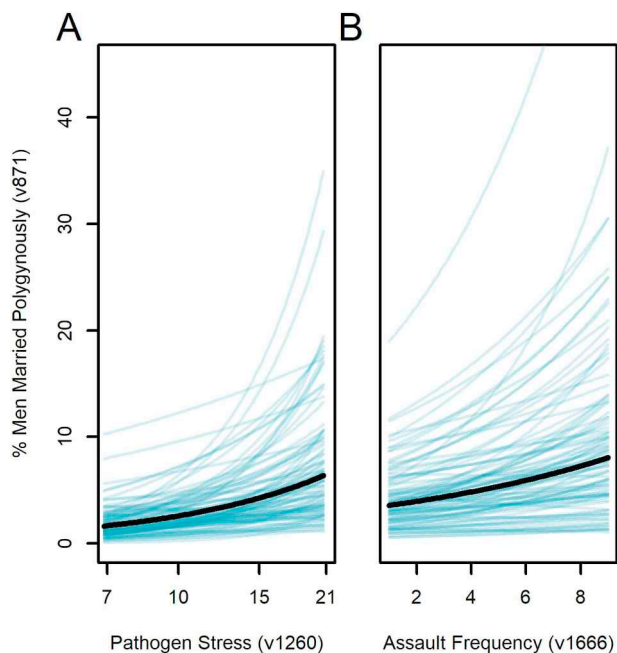


Fig. 5. Illustrating the predicted association between percent of married men who are polygynous (v871) and A) pathogen stress and B) assault frequency. Graphs plot the mean prediction (thick line) and 100 samples randomly drawn from the posterior to illustrate the degree of uncertainty.

4.1. Phylogeny and ecology explain behavioral variation

We found that the risk for both type I and type II errors may be elevated when ignoring phylogeny, consistent with simulation studies (Harvey & Rambaut, 1998; Lindfors et al., 2010; Nunn, 2011) and previous cross-cultural analyses contrasting phylogenetic or autocorrelation methods with standard statistical tests (e.g. Bergerhoff Mulder et al., 2001; Dow, 1993). The false positive rate found here (4/26 tests, 15%) is substantially lower than that in the latter two studies (which were around 50%), perhaps partly because the SCCS populations are indeed more independent than other samples (though see Dow & Eff, 2008; Dow, 1993), but perhaps more so because our analyses included multiple predictors rather than conducting bivariate tests, thus reducing the number of significant associations to begin with.

Our results also appear to support previous studies contrasting current ecology and population history as predictors of human behavioral variation in smaller geographic areas (Bergerhoff Mulder et al., 2001; Guglielmino et al., 1995; Hewlett et al., 2002; Mathew & Perreault, 2015), typically finding stronger evidence for the latter, as we did on a global scale. However, we again emphasize that this pattern is consistent with various processes (genetic or cultural inheritance, niche conservatism, unmeasured ecological similarity); for instance, other aspects of marriage, such as bridewealth or divorce, also show strong phylogenetic signal in East Africa, yet this may be partly due to closely related groups inhabiting similar environments (e.g. Nilotic and Cushitic groups live in drier areas than Bantu; Bergerhoff Mulder et al., 2001). Likewise, adding more predictors would have increased the proportion of variance explained by socio-ecology, yet many aspects of

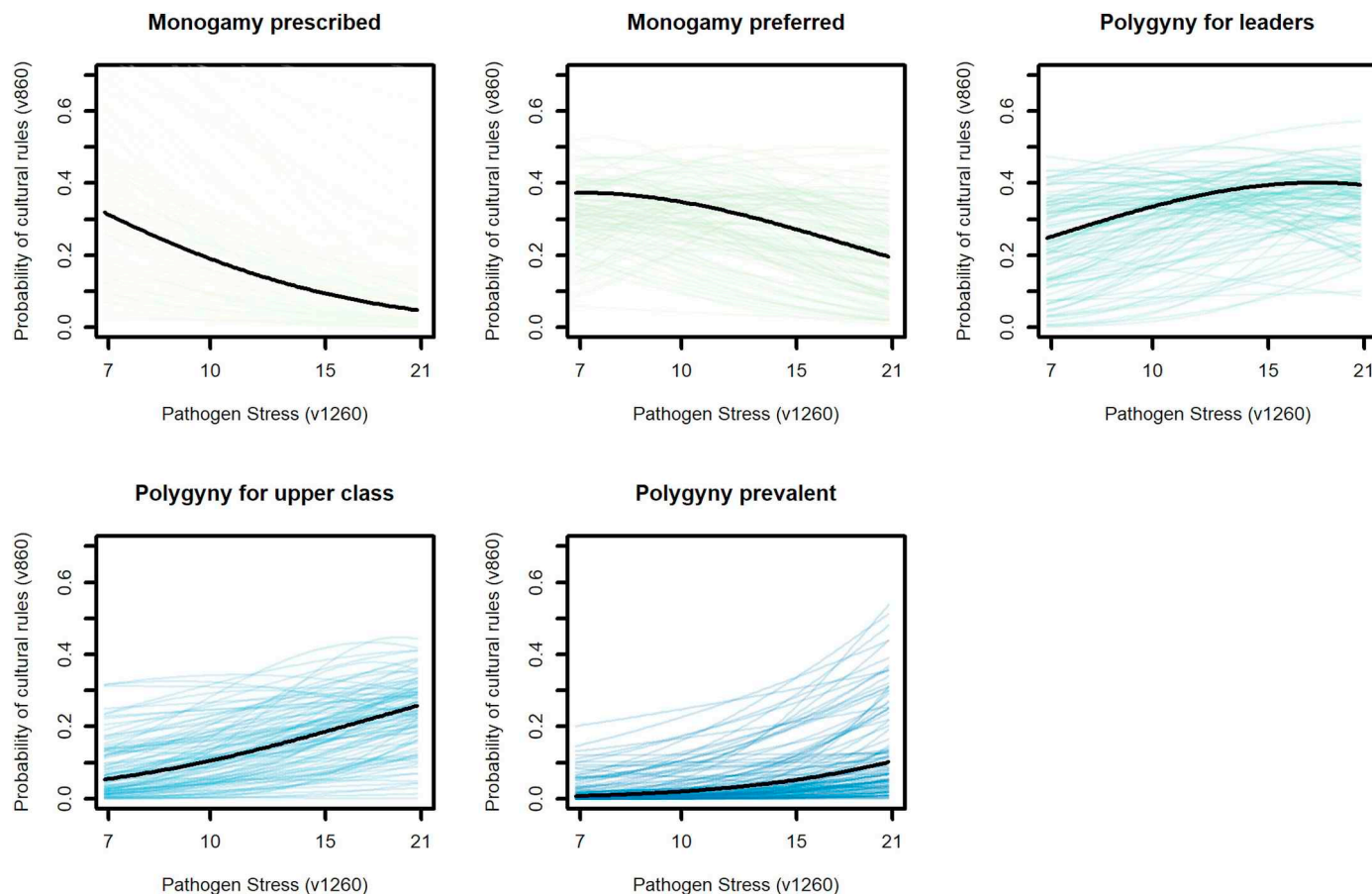


Fig. 6. Illustrating the predicted association between cultural rules constraining polygyny (v860) and pathogen stress (see Fig. S1 for assault frequency). Graphs plot the mean prediction (thick black line) and 100 samples randomly drawn from the posterior to illustrate the degree of uncertainty.

current socio-ecology have themselves been shaped by cultural evolution; e.g. being Bantu accounts for 50% of the variance in female contribution to agriculture (Burton & White, 1984). Thus, it may be impossible to partition variance in behavior at the societal level into current ecology and cultural history, just like partitioning transmission into vertical and horizontal- (Borgerhoff Mulder et al., 2006; Towner et al., 2015), or individual behavior into genetic and environmental components (Joseph, 2014) is difficult. We therefore prefer to highlight the value of combining population history and current socio-ecological conditions in comparative studies (see also Section 4.3.).

In sum, we provide strong evidence that accounting for phylogeny is crucial in cross-cultural studies, as emphasized by previous authors (Borgerhoff Mulder et al., 2001; Mace & Pagel, 1994; Nunn, 2011), because it (i) prevents statistical errors arising from the non-independence of data points, and (ii) provides a more comprehensive explanation of behavioral variation. We therefore urge future SCCS studies to make use of the present supertree, and hope to soon expand this approach to larger cross-cultural samples such as the Ethnographic Atlas. Possible expansions of the phylogenetic methods used here are models that distinguish between the origins and maintenance of traits and allow selective regimes to vary by clade (e.g. Ross et al., 2016), or account for multiple sources of covariance such as timing of the ethnographic present and geography, in addition to phylogeny (e.g. Ringen & Jaeggi, 2018).

4.2. Synthesis and caveats in understanding marriage patterns

Previous work on marriage patterns in the SCCS relied on testing a small number of predictors in varying subsets of the sample due to missing data. Here we were able to include all previously considered predictors of polygyny by imputing missing values. Our model thus provides the most comprehensive and stringent analysis of polygyny in the SCCS to date, with associations between predictors and polygyny being contingent on all other variables in the model (thus reducing the risk of spurious correlations), in addition to controlling for phylogenetic non-independence (thus reducing the risk of false positives). However, such associations are subject to various caveats as discussed below.

The two strongest predictors of polygyny, pathogen stress and assault frequency, are consistent with aspects of sexual selection theory (Table 1), specifically with female choice of good genes (*model 1b*) and male intra-sexual competition and coercion (*model 2*). Increased pathogen stress has been argued to exacerbate variation in male genetic quality (Low, 1988, 1990), thus fewer men are able to develop healthy phenotypes. This arguably lowers the polygyny threshold as more women should choose men with good genes (a non-rival form of male wealth) to increase the expected fitness of their offspring. Assault frequency could be related to both male intra-sexual competition and male coercion because perpetrators and victims of assault are not specified (Marlowe, 2003), however, these tend to overwhelmingly be both male cross-culturally (Daly & Wilson, 1988; Fry & Söderberg, 2013), suggesting that assault frequency mostly captures male-male competition. Indeed, another recent SCCS study found an association between polygyny and male aggressiveness (Carter & Kushnick, 2018), providing further support for harem-defense polygyny (*model 2*). In addition, high assault frequency in a population could put a premium on male protection of women and their offspring, another non-rival form of wealth, and could thus lead to polygyny based on female choice (Hooper, Gurven, & Kaplan, 2014). Lastly, assault frequency could also indirectly influence polygyny by altering the sex ratio, though sex ratios in small-scale societies can be quite stochastic (Kramer, Schacht, & Bell, 2017). Note though that male-male competition does not need to take violent forms (Schacht, Rauch, & Borgerhoff Mulder, 2014), and that we cannot here determine whether assault frequency is a cause or a consequence of greater polygyny.

While pathogen stress and assault frequency are thus robustly associated with polygyny in this sample, we should also emphasize that

interpreting population-level associations as evidence for individual-level behavior can lead to erroneous conclusions (i.e. the ecological fallacy, or ‘Simpson’s paradox’); for instance, negative associations between polygyny and child health outcomes at the regional level disappear or even turn positive when using individual-level data within ethnic groups (Lawson et al., 2015). Similarly, population-level associations between parasite prevalence and various attitudes and behaviors were not supported when using individual data and appropriate multilevel modeling techniques (Ross & Winterhalder, 2016). Lastly, covariation between pathogen stress (or any other predictor) and other environmental features could generate spurious associations (Bromham, Hua, Cardillo, Schneemann, & Greenhill, 2018), though we cannot think of a third variable that would correlate with both pathogens and polygyny but was not considered in our model. Thus, the associations of pathogen stress and assault frequency with polygyny come with several caveats, as is typical for the often crude comparative record. Nonetheless, our results lend credence to previous findings that did not control for phylogeny or most of the other covariates in our model (e.g. Low, 1988, 1990; Marlowe, 2003), making them subject to valid skepticism (Bromham et al., 2018).

Predictors capturing variance in male resources (stratification, wealth inequality), i.e. resource-defense polygyny (*model 1a*), were inconsistently associated with polygyny despite strong evidence for a polygyny threshold based on male wealth in some populations (Borgerhoff Mulder, 1990). Nonetheless, the highest level of wealth inequality was expected to result in more polygyny than the lowest level with high probability (v871: PP = 0.95. v860: PP = 0.99), despite the large amount of missing data and associated uncertainty in this variable. By contrast, the highest level of stratification was strongly associated with reduced polygyny (v871: PP = 0.98. v860: PP = 0.99). In combination, these two variables arguably recreate the inverse-U shaped association between polygyny and wealth inequality found in a recent cross-cultural study, driven by increasing levels of polygyny with inequality among foragers, horticulturalists and pastoralists, but low levels of polygyny in highly unequal agricultural societies (Ross et al., 2018) – note that this contrasts with other measures such as maximum harem size or variance in reproductive success, which increase with stratification (Betzig, 1982; Betzig, 2012), while the correlation between male status and reproductive success was unaffected by it (von Rueden & Jaeggi, 2016). In this context, it is worth considering that wealth inequality was only coded for foragers (Pryor, 2003) and imputed for all other populations, hence the influence of this variable may only hold in a range of societies with relatively low levels of wealth inequality to begin with. Conversely, stratification captures the whole breadth of human societal complexity, from egalitarian foragers to despotic states, with their low levels of polygyny (Ross et al., 2018). Our results do not allow us to further comment on the reasons for the low levels of polygyny in agricultural societies (see Fortunato & Archetti, 2010; Fortunato, 2011; Henrich, Boyd, & Richerson, 2012; Oh et al., 2016; Ross et al., 2018; hypotheses 4 and 5 in Table 1), though it is interesting that both the cultural norms constraining polygyny (as postulated by Henrich et al., 2012) and the actual percentage of men married polygynously decline with societal complexity. Lastly, greater female contribution to agriculture was reliably associated with at least one measure of polygyny (v871: PP = 0.96), and that more marriages were polygynous when female choice is potentially constrained by arranged marriages (v871: PP = 0.94. v860: PP = 0.93). Note that some predictors previously found to be associated with polygyny such as warfare (Ember et al., 2007) received little support in our model, which could suggest a true statistical artefact revealed once other predictors were included and phylogeny accounted for, but could also be due to slightly different measures of polygyny and warfare used.

In sum, our analysis provides the strongest support for a polygyny threshold based on male genetic quality and for harem-defense polygyny (*models 1b* and *2* in Table 1), and weaker support for a polygyny threshold based on male wealth (*model 1a*), as potentially curtailed by

the importance of rival wealth (*model 4*) and/or socially imposed norms of monogamy (*model 5*). We found no support for marriage patterns to be influenced by the sex ratio (*model 3*), despite this prediction stemming from one of the most recent formulations of sexual selection theory (Kokko & Jennions, 2008) and receiving good support in some studies (Pollet & Nettle, 2009; Schacht & Borgerhoff Mulder, 2015). This could be due to the poor quality of the SCCS sex ratio variable, or because sex ratio effects operate through mechanisms better captured by other variables in the model.

4.3. Conclusion

What does it mean to combine the influence of population history and socio-ecology on marriage patterns? Marriage patterns are a population-level trait that arises from the decisions of various individuals (bride and groom, their families and communities) as they respond to incentive structures in their environment as well as the inherited social norms of their culture (Henrich et al., 2012; Walker et al., 2011); indeed, one of our measures of polygyny (v860) literally consists of cultural norms, whereas the other (v871) arguably more directly reflects individual behavior. In this sense marriage goes beyond mere mating, which should only strengthen the influence of cultural history. As mentioned above, the influence of culture and ecology can be difficult to disentangle as populations select similar ecological niches as their ancestors (e.g. Bantu speakers in East Africa choosing wetter areas; Borgerhoff Mulder et al., 2001), which influences their behavior (e.g. 50% of the variation in female contribution to subsistence explained by being Bantu; Burton & White, 1984) in ways that affect marriage patterns (e.g. more men married polygynously in populations with greater female contribution to agriculture; see Fig. 3a). Even if culture and ecology were perfectly separable, it is reasonable to assume that both should influence marriage patterns. Cultural evolution theory emphasizes that virtually any norm can be stabilized by punishment (Boyd & Richerson, 1992), and marriage patterns are subject to various norms (such as rules of exogamy or wealth inheritance, incest taboos, etc., see Henrich et al., 2012), and that therefore multiple stable strategies may exist in a given ecology (Boyd & Richerson, 1990); in other words, there are several plausible ways in which a society can function. Current norms may thus be somewhat arbitrary, slow to change, and path-dependent, i.e. they reflect historical practices as people socially learn the norms of their culture (Richerson & Boyd, 2004), as evidenced here by the strong influence of phylogeny (Fig. 4). On the other hand, it is undeniable that marriage patterns do reflect local fitness incentives, in ways that are often consistent with mating patterns in other species (see studies summarized in Table 1, results presented here). To the extent that both measures of polygyny are equally associated with a predictor, it would then appear that cultural norms and individual behavior have converged on locally optimal solutions; for instance, monogamy prevails in both norms (v860) and actual marriages (v871) in the most complex societies (see Fig. 3). While comparative analyses such as ours can thus statistically integrate predictions from cultural evolution and behavioral ecology, more detailed studies are needed to uncover the extent to which individual behaviors reflect inherited norms or current incentive structures (e.g. Daly & Wilson, 2010; Du & Mace, 2018), or whether and how norms are enforced (e.g. Mathew & Boyd, 2011).

In summary, we used a new supertree of human populations and modern statistical methods to provide one of the most comprehensive cross-cultural analyses of marriage patterns to date. We found that (i) ignoring phylogeny increases the risk of both false positives and false negatives, (ii) phylogeny accounts for a substantial proportion of human behavioral variation, highlighting the potential for comparative studies to combine population history and current socio-ecology, and (iii) the strongest predictors of polygyny, pathogen stress and assault frequency, are consistent with sexual selection accounts of human marriage systems. In the future, the present supertree and analysis approach can be applied to any number of traits coded in the SCCS, thus

expanding the phylogenetic study of human behavioral variation to a global scale.

4.4. Open practices/data availability

The data, phylogenetic tree, and all R code associated with this research are available at <https://github.com/rianaminocher/polygyny-sccs>.

Acknowledgements

Monique Borgerhoff Mulder and David W. Lawson provided extensive comments on earlier drafts that greatly improved this manuscript. We also thank Erik Ringen and Jordan Martin for methodological feedback. Any remaining errors are solely the authors'. P.D. was supported by the Czech Science Foundation (GACR) Grant #18-23889S. This study was not preregistered.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.evolhumbehav.2018.11.003>.

References

- Barry, H., III, & Paxson, L. M. (1971). Infancy and early childhood: Cross-cultural codes 2. *Ethnology*, 10(4), 466.
- Barry, H., III, & Schlegel, A. (1982). Cross-cultural codes on contributions by women to subsistence. *Ethnology*, 21, 165–188.
- Betzig, L. (2012). Means, variances, and ranges in reproductive success: Comparative evidence. *Evolution and Human Behavior*, 33(4), 309–317.
- Betzig, L. L. (1982). Despotism and differential reproduction: A cross-cultural correlation of conflict asymmetry, hierarchy, and degree of polygyny. *Ethology and Sociobiology*, 3(4), 209–221.
- Bininda-Emonds, O. R. P. (2014). An introduction to supertree construction (and partitioned phylogenetic analyses) with a view toward the distinction between gene trees and species trees. In L. Z. Garamszegi (Ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology* (pp. 49–76). Heidelberg: Springer.
- Bininda-Emonds, O. R. P., Cardillo, M., Jones, K. E., MacPhee, R. D. E., Beck, R. M. D., Grenyer, R., ... Purvis, A. (2007). The delayed rise of present-day mammals. *Nature*, 446(7135), 507–512.
- Borgerhoff Mulder, M. (1988). The relevance of the polygyny threshold model for humans. In C. Mascie-Taylor, & A. Boyce (Eds.), *Mating patterns* (pp. 209–230). Cambridge: Cambridge University Press.
- Borgerhoff Mulder, M. (1990). Kipsigis women's preferences for wealthy men: Evidence for female choice in mammals? *Behavioral Ecology and Sociobiology*, 27(4), 255–264.
- Borgerhoff Mulder, M., George-Cramer, M., Eshleman, J., & Ortolani, A. (2001). A study of East African kinship and marriage using a phylogenetically based comparative method. *American Anthropologist*, 103(4), 1059–1082.
- Borgerhoff Mulder, M., Nunn, C. L., & Towner, M. C. (2006). Cultural macroevolution and the transmission of traits. *Evolutionary Anthropology*, 15(2), 52–64.
- Boyd, R., & Richerson, P. J. (1990). Group selection among alternative evolutionarily stable strategies. *Journal of Theoretical Biology*, 145, 331–342.
- Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, 13(3), 171–195.
- Bromham, L., Hua, X., Cardillo, M., Schneemann, H., & Greenhill, S. J. (2018). Parasites and politics: Why cross-cultural studies must control for relatedness, proximity and covariation. *Royal Society Open Science*, 5, 181100.
- Broude, G. J., & Greene, S. J. (1983). Cross-cultural codes on husband-wife relationships. *Ethnology*, 22(3), 263.
- Burton, M. L., & White, D. R. (1984). Sexual division of labor in agriculture. *American Anthropologist*, 86(3), 568–583.
- Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Multivariate imputation by chained equations. *Journal of Statistical Software*, 45(3), 1–67.
- Carter, T.-L., & Kushnick, G. (2018). Male aggressiveness as intrasexual contest competition in 78 societies. *Behavioral Ecology and Sociobiology*, 72, 83.
- Chaudhary, N., Salali, G. D., Thompson, J., Dyble, M., Page, A., Smith, D., ... Migliano, A. B. (2015). Polygyny without wealth: Popularity in gift games predicts polygyny in BaYaka Pygmies. *Royal Society Open Science*, 2(5), 150054.
- Chisholm, J. S., & Burbank, V. K. (1991). Monogamy and polygyny in Southeast Arnhem land: Male coercion and female choice. *Ethology and Sociobiology*, 12(4), 291–313.
- Clutton-Brock, T., & Vincent, A. (1991). Sexual selection and the potential reproductive rates of males and females. *Nature*, 351, 58–60.
- Cowlshaw, G., & Mace, R. (1996). Cross-cultural patterns of marriage and inheritance: A phylogenetic approach. *Ethology and Sociobiology*, 17, 87–97.
- Daly, M., & Wilson, M. (1988). *Homicide*. Piscataway NJ: Transaction Publishers.
- Daly, M., & Wilson, M. (2010). Cultural inertia, economic incentives, and the persistence of Southern violence. In M. Schaller, A. Norenzayan, S. Heine, T. Yamagishi, & T.

- Kameda (Eds.). *Evolution, culture, and the human mind* (pp. 229–241). New York: Psychology Press.
- Development Core Team, R. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Dow, M. M. (1993). Saving the theory: On Chi-square tests with cross-cultural survey data. *Cross-Cultural Research*, 27(3–4), 247–276.
- Dow, M. M., & Eff, E. A. (2008). Global, regional, and local network autocorrelation in the Standard Cross-Cultural Sample. *Cross-Cultural Research*, 42(2), 148–171.
- Dow, M. M., & Eff, E. A. (2009). Multiple imputation of missing data in cross-cultural samples. *Cross-Cultural Research*, 43(3), 206–229.
- Du, J., & Mace, R. (2018). Parental investment in Tibetan populations does not reflect stated cultural norms. *Behavioral Ecology*, 29(January), 106–116.
- Duda, P., & Zrzavý, J. (2016). Human population history revealed by a supertree approach. *Scientific Reports*, 6, 29890.
- Duda P. and Zrzavý J., Towards a global phylogeny of human populations based on genetic and linguistic data. In: Sahle Y., Bentz C. and Reyes-Centeno H., (Eds.), *Modern Human Origins and Dispersal*, in press, Kerns Verlag; Tübingen.
- Ember, C. R., & Ember, M. (1992). Warfare, aggression, and resource problems: Cross-cultural codes. *Behavior Science Research*, 26(1–4), 169–226.
- Ember, M. (1974). Warfare, sex ratio, and polygyny. *Ethnology*, 13(2), 197–206.
- Ember, M., Ember, C. R., & Low, B. S. (2007). Comparing explanations of polygyny. *Cross-Cultural Research*, 41(4), 428–440.
- Emlen, S., & Oring, L. (1977). Ecology, sexual selection, and the evolution of mating systems. *Science*, 197(4300), 215–223.
- Felsenstein, J. (1985). Phylogenies and the comparative method. *The American Naturalist*, 125(1), 1–15.
- Fortunato, L. (2011). Reconstructing the history of marriage strategies in indo-European-speaking societies: Monogamy and polygyny. *Human Biology*, 83(1), 87–105.
- Fortunato, L., & Archetti, M. (2010). Evolution of monogamous marriage by maximization of inclusive fitness. *Journal of Evolutionary Biology*, 23(1), 149–156.
- Fortunato, L., Holden, C., & Mace, R. (2006). From bridewealth to dowry? *Human Nature*, 17(4), 355–376.
- Fry, D. P., & Söderberg, P. (2013). Lethal aggression in mobile forager bands and implications for the origins of war. *Science*, 341(6143), 270–273.
- Guglielmino, C. R., Viganotti, C., Hewlett, B., & Cavalli-Sforza, L. L. (1995). Cultural variation in Africa - Role of mechanisms of transmission and adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, 92(16), 7585–7589.
- Hadfield, J. (2016). *MCMCglmm Course notes. 144*. Retrieved from <http://cran.us-r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf>.
- Hadfield, J. D. (2010). MCMC methods for multi-response generalised linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, 33(2), 1–22.
- Hadfield, J. D., & Nakagawa, S. (2010). General quantitative genetic methods for comparative biology: Phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *Journal of Evolutionary Biology*, 23(3), 494–508.
- Hansen, T. F. (2014). Use and misuse of comparative methods in the study of adaptation. In L. Z. Garamszegi (Ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology* (pp. 351–380). Heidelberg: Springer.
- Hartung, J. (1982). Polygyny and inheritance of wealth. *Current Anthropology*, 23(1), 1–12.
- Harvey, P. H., & Rambaut, A. (1998). Phylogenetic extinction rates and comparative methodology. *Proceedings of the Royal Society B: Biological Sciences*, 265(May), 1691–1696.
- Henrich, J., Boyd, R., & Richerson, P. J. (2012). The puzzle of monogamous marriage. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1589), 657–669.
- Hewlett, B. S., De Silvestri, A., & Guglielmino, C. R. (2002). Semes and genes in Africa. *Current Anthropology*, 43(2), 313–321.
- Hooper, P. L., Gurven, M., & Kaplan, H. S. (2014). Social and economic underpinnings of human biodemography. In M. Weinstein, & M. A. Lane (Eds.), *Sociality, hierarchy, health: Comparative biodemography* (pp. 169–195). Washington, DC: The National Academies Press.
- Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature*, 1–5.
- Jordan, F. M., Gray, R. D., Greenhill, S. J., & Mace, R. (2009). Matrilineal residence is ancestral in Austronesian societies. *Proceedings of the Royal Society B - Biological Sciences*, 276(1664), 1957–1964.
- Joseph, J. (2014). *The trouble with twin studies: A reassessment of twin research in the social and behavioral sciences*. New York, NY: Routledge.
- Kokko, H., & Jennions, M. D. (2008). Parental investment, sexual selection and sex ratios. *Journal of Evolutionary Biology*, 21(4), 919–948.
- Kramer, K. L., Schacht, R., & Bell, A. (2017). Adult sex ratios and partner scarcity among hunter-gatherers: Implications for dispersal patterns and the evolution of human sociality. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 372(1729), 20160316.
- Lawson, D., James, S., Ngadaya, E., Ngowi, B., Mfinanga, S. G. M., & Borgerhoff Borgerhoff, M. (2015). No evidence that polygynous marriage is a harmful cultural practice in northern Tanzania. *Proceedings of the National Academy of Sciences of the United States of America*, 1–6.
- Lindenfors, P., Revell, L. J., & Nunn, C. L. (2010). Sexual dimorphism in primate aerobic capacity: A phylogenetic test. *Journal of Evolutionary Biology*, 23(6), 1183–1194.
- Low, B. S. (1988). Pathogen stress and polygyny in humans. In L. L. Betzig, M. Borgerhoff Mulder, & P. Turke (Eds.), *Human reproductive behavior: A Darwinian perspective* (pp. 115–127). Cambridge: Cambridge University Press.
- Low, B. S. (1990). Marriage systems and pathogen stress in human societies. *American Zoologist*, 30(2), 325–340.
- Lukas, D., & Clutton-Brock, T. H. (2013). The evolution of social monogamy in mammals. *Science*, 526.
- Mace, R., & Pagel, M. (1994). The comparative method in anthropology. *Current Anthropology*, 35(5), 549–564.
- Mace, R., & Sear, R. (2005). Are humans cooperative breeders? In E. Voland, A. Chasiotis, & W. Schiefelhövel (Eds.), *Grandmotherhood: The evolutionary significance of the second half of female life* (pp. 143–159). Rutgers: The State University.
- Marlowe, F. (2000). Paternal investment and the human mating system. *Behavioural Processes*, 51(1–3), 45–61.
- Marlowe, F. W. (2003). The mating system of foragers in the Standard Cross-Cultural Sample. *Cross-Cultural Research*, 37(3), 282–306.
- Mathew, S., & Boyd, R. (2011). Punishment sustains large-scale cooperation in prestate warfare. *Proceedings of the National Academy of Sciences*, 108(28), 11375–11380.
- Mathew, S., & Perreault, C. (2015). Behavioural variation in 172 small-scale societies indicates that social learning is the main mode of human adaptation. *Proceedings of the Royal Society B - Biological Sciences*, 282, 20150061.
- McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Boca Raton, FL: CRC Press.
- Moravec, J. C., Atkinson, Q., Bowern, C., Simon, J., Jordan, F. M., Ross, R. M., ... Marsland, S. (2018). Post-marital residence patterns show lineage-specific evolution. *Evolution and Human Behavior*, 39(6), 594–601.
- Murdock, G. P. (1967). Ethnographic atlas: A summary. *Ethnology*, 6(2), 109–236.
- Murdock, G. P., & Provost, C. (1973). Measurement of cultural complexity. *Ethnology*, 12(4), 379.
- Murdock, G. P., & White, D. R. (1969). Standard Cross-Cultural Sample. *Ethnology*, 8(4), 329–369.
- Murdock, G. P., & Wilson, S. F. (1972). Settlement patterns and community organization: Cross-cultural codes 3. *Ethnology*, 11(3), 254.
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2), 133–142.
- Nettle, D., Gibson, M. A., Lawson, D. W., & Sear, R. (2013). Human behavioral ecology: Current research and future prospects. *Behavioral Ecology*, 24(5), 1031–1040.
- Nunn, C. L. (2011). *The comparative approach in evolutionary anthropology and biology*. University of Chicago Press.
- Oh, S.-Y., Ross, C. T., Borgerhoff Mulder, M., & Bowles, S. (2016). *The decline in polygyny: An interpretation*. (Santa Fe).
- Opie, C., Shultz, S., Atkinson, Q. D., Currie, T., & Mace, R. (2014). Phylogenetic reconstruction of Bantu kinship challenges Main Sequence Theory of human social evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 111(49), 17414–17419.
- Osmond, M. W. (1965). Toward monogamy: A cross-cultural study of correlates of type of marriage. *Social Forces*, 44, 8–16.
- Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, 401(6756), 877–884.
- Pagel, M., Atkinson, Q. D., S. Calude, A., & Meade, A. (2013). Ultraconserved words point to deep language ancestry across Eurasia. *Proceedings of the National Academy of Sciences*, 110(21), 8471–8476.
- Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). Output analysis and diagnostics for MCMC. *R News*, 6(1), 7–11.
- Pollet, T. V., & Nettle, D. (2009). Market forces affect patterns of polygyny in Uganda. *Proceedings of the National Academy of Sciences*, 106(7), 2114–2117.
- Pryor, F. L. (2003). Economic systems of foragers. *Cross-Cultural Research*, 37(4), 393–426.
- Quinlan, R. J., & Quinlan, M. B. (2007). Evolutionary ecology of human pair-bonds: Cross-cultural tests of alternative hypotheses. *Cross-Cultural Research*, 41, 149–169.
- Richerson, P. J., & Boyd, R. (2004). *Not by genes alone: How culture transformed human evolution*. University of Chicago Press.
- Ringen, E. J., & Jaeggi, A. V. (2018). Daily food sharing in nonindustrial societies: Effects of subsistence ecology, food storage technology, and spatial/phylogenetic distance. *American Journal of Physical Anthropology*, 165, 226.
- Ross, C. T., Mulder, M. B., Oh, S., Bowles, S., Beheim, B., Bunce, J., ... Ziker, J. (2018). Greater wealth inequality, less polygyny: Rethinking the polygyny threshold model. *Journal of the Royal Society Interface*, 15, 20180035.
- Ross, C. T., Strimling, P., Ericksen, K. P., Lindenfors, P., & Borgerhoff Mulder, M. (2016). The origins and maintenance of female genital modification across Africa: Bayesian phylogenetic modeling of cultural evolution under the influence of selection. *Human Nature*, 27(2), 173–200.
- Ross, C. T., & Winterhalder, B. (2016). A hierarchical bayesian analysis of parasite prevalence and sociocultural outcomes: The role of structural racism and sanitation infrastructure. *American Journal of Human Biology*, 28(1), 74–89.
- von Rueden, C. R., & Jaeggi, A. V. (2016). Men's status and reproductive success in 33 nonindustrial societies: Effects of subsistence, marriage system, and reproductive strategy. *Proceedings of the National Academy of Sciences*, 113(39), 10824–10829.
- Sanderson, S. K. (2001). Explaining monogamy and polygyny in human societies: Comment on Kanazawa and still. *Social Forces*, 80(1), 329–335.
- Schacht, R., & Borgerhoff Mulder, M. (2015). Sex ratio effects on reproductive strategies in humans. *Royal Society Open Science*, 2(1), 140402.
- Schacht, R., Rauch, K. L., & Borgerhoff Mulder, M. (2014). Too many men: The violence problem? *Trends in Ecology & Evolution*, 1–9.
- Sellen, D. W., & Hruschka, D. J. (2004). Extracted-food resource-defense polygyny in native Western north American societies at contact. *Current Anthropology*, 45(5), 707–714.
- Sheehan, O., Watts, J., Gray, R. D., & Atkinson, Q. D. (2018). The coevolution of intensive agriculture and socio-political hierarchy. *Proceedings of the National Academy of Sciences*, (11), 1–6.
- Shultz, S., Opie, C., & Atkinson, Q. D. (2011). Stepwise evolution of stable sociality in

- primates. *Nature*, 479(7372), 219–222.
- Thierry, B. (2008). Primate socioecology, the lost dream of ecological determinism. *Evolutionary Anthropology*, 17(2), 93–96.
- Towner, M. C., Grote, M., & Borgerhoff Mulder, M. (2015). Problems modelling behavioural variation across Western north American Indian societies. *Proceedings of the Royal Society B - Biological Sciences*, 283, 20152184.
- Trivers, R. (1972). In B. Campbell (Ed.), *Parental investment and sexual selection* (pp. 136–179). Chicago, IL: Aldine Sexual selection and the descent of man.
- de Villemereuil, P., & Nakagawa, S. (2014). General quantitative genetic methods for comparative biology. In L. Z. Garamszegi (Ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology* (pp. 287–303). Berlin: Springer.
- Walker, R. S., Hill, K. R., Flinn, M. V., & Ellsworth, R. M. (2011). Evolutionary history of hunter-gatherer marriage practices. *PLoS One*, 6(4), e19066.
- White, D. R. (1988). Rethinking polygyny: Co-wives, codes, and cultural systems. *Current Anthropology*, 29(4), 529–572.
- White, D. R., & Burton, M. L. (1988). Causes of polygyny: Ecology, economy, kinship, and warfare. *American Anthropologist*, 90(4), 871–887.
- White, D. R., Whiting, J. W. M., & Burton, M. L. (1986). Climate and subsistence codes. *World Cultures*, 2(2).
- Whiting, J. W. M., Sodergren, J. A., & Stigler, S. M. (1982). Winter temperature as a constraint to the migration of preindustrial peoples. *American Anthropologist*, 84(2), 279–298.
- Whyte, M. K. (1978). Cross-cultural codes dealing with the relative status of women. *Ethnology*, 17(2), 211.
- Whyte, M. K. (1985). The status of women in preindustrial societies: Dependent and independent variables. *World Cultures*, (4), 1.
- Winterhalder, B., & Smith, E. A. (2000). Analyzing adaptive strategies: Human behavioral ecology at twenty-five. *Evolutionary Anthropology*, 9, 51–72.
- Zhou, X., & Reiter, J. P. (2010). A note on Bayesian inference after multiple imputation. *The American Statistician*, 64(1987), 159–163.

Phylogenetic analysis of polygyny in SCCS – Supplementary Informations

Contents

Selecting predictor variables.....	2
Group 1: Subsistence	3
Group 2: Violence	4
Group 3: Stratification	5
Group 4: Climate	6
Independent predictors	7
Table S1: Summary of included variables	8
Table S2: Summary of phylogenetic and non-phylogenetic models	10
Figure S1.....	11
Supertree inference and time-calibration	12
References	14

Selecting predictor variables

We considered virtually all variables that have previously been tested as predictors of polygyny in the SCCS. Two of these variables, "mort" and "plow" were provided by Carol Ember; the rest were available as part of the publicly available SCCS. In order to reduce the number of covariates to include in our models, we employed the following procedure:

- 1) Attempt to qualitatively reproduce previously published analyses using the same variables and statistical methods. This was intended to check that the variables we had access to were still comparable to the originally used ones, since the SCCS is updated regularly
- 2) Group variables by the relevant socio-ecological conditions they are meant to capture, i.e. subsistence, violence, stratification, and climate, and eliminate redundant variables. Then, within each group:
 - a. Use variance inflation factors (VIF's) to eliminate highly collinear predictors, such that all VIF's < 3
 - b. Run a principal components analysis (PCA) on the remaining predictors, to assess which ones capture most of the variance in the data. Select individual variables if they load highly on a component that explains most of the variance. Note that we had to use the original variables rather than the principal components in our eventual models because the latter would not be comparable across each imputed dataset. In practice, this made little difference as the relevant components were virtually identical with a single variable (see below for details)

Step 1 eliminated several variables because we were unable to qualitatively reproduce the original results. These included variables such as fraternal interest groups, small island, warfare for plunder, and marriage of captives (most of which used by White & Burton, 1988). Thus, we were left with a total of 27 potential predictors that entered step 2.

We then grouped these 27 variables into subsistence, violence, stratification, and climate groups. We decided to include population density, father-infant proximity, pathogen stress, sex ratio, and arranged female marriages as independent predictors because these variables did not fit into the other groups.

A number of these variables measure the exact same ecological feature, namely different measures of dependence on hunting, fishing or gathering, subsistence mode, and female contribution to subsistence. In these cases we chose to include the variable that had fewer missing data. Thus, five variables were eliminated because of redundancy, all in the subsistence group (see below for details). The elimination of collinear variables using VIF's and PCA's within each group (details given below) plus the independent predictors resulted in the eleven predictors we ultimately included in our models.

All R code and data for the variable selection will available at [\[link\]](#).

Group 1: Subsistence

variable	description	N	type
v7	Fishing- Contribution to Food Supply	184	ordinal
v9	Hunting- Contribution to Food Supply	182	ordinal
v11	Gathering- Contribution to Food Supply	182	ordinal
v205	Dependence on Fishing (Atlas 3)	186	ordinal
v204	Dependence on Hunting (Atlas 3)	186	ordinal
v203	Dependence on Gathering (Atlas 1)	186	ordinal
v885	Female Contribution to Subsistence: Ethnographic Atlas	152	ordinal
v890	Female Contribution to Subsistence: Average of Three Scores (Ordinated)	185	ordinal
v833	Subsistence Economy: Dominant Mode	186	categorical
v858	Subsistence Type - Ecological Classification D. White, 1984, after Karen and Jeffrey Paige (L981)	186	categorical
v586	Relative Time and Effort Expended on Subsistence Activities	88	ordinal
plow	Presence of plow (Ember)	186	binomial

The above table shows all candidate variables in the subsistence group. Of these, we removed variables v7, v9, and v11 in favor of v205, v204, v203, which appear to measure the same ecological variables, because of fewer missing data. We removed v885 in favor of v890 because of fewer missing data, and given that the variable is a composite of three different measures of female contribution to subsistence. We used v858 over v833 because the variable contained finer-scale data (11 levels compared with 8 levels). To express mode of subsistence better by type we re-leveled the variable v858 into 5 categories: 1) forager = gathering, hunting and/or marine animals, 2) fishing; complex forager = anadromous fishing, mounted hunting; 3) pastoralist = pastoralism; 4) horticulturalist = shifting cultivation with digging sticks or wooden hoes, shifting cultivation with metal hoes, horticultural gardens or tree fruits; 5) intensive agriculture = intensive agriculture with no plow, intensive agriculture with plow.

Assessing the VIFs for these predictors, we removed v858.

```
##          GVIF Df  GVIF^(1/(2*Df))
## v586  1.307163  1      1.143312
## v858  7.745689  4      1.291613
## v890  1.574288  1      1.254706
## v204  2.774452  1      1.665669
## v205  1.789366  1      1.337672
## v203  3.555771  1      1.885675
## plow  2.876114  1      1.695911
```

We performed a PCA on time in subsistence (v586), female contribution to agriculture (v890), dependence on gathering (v203), -hunting (v204), -fishing (v205), and the presence of the plow, with N=88 populations having data on all these variables. The PCA showed that a single component explained 97.99% of the variance across populations, and this component almost exclusively consisted of v890 (loading = 0.9997, all other variable loadings < |0.02|). Thus, we selected v890 for our analysis, to represent variation in mode of subsistence.

```
## Standard deviations (1, .., p=6):
## [1] 14.1099918  1.4732226  1.0706526  0.6194484  0.5425225  0.3128088
##
## Rotation (n x k) = (6 x 6):
##           PC1          PC2          PC3          PC4
## v586  0.015226776  0.1205050  0.05201853  0.425866083
## v890  0.999735116 -0.0104369 -0.01309271 -0.004040141
## v204  0.005796062  0.7082669  0.14612123  0.568043370
## v205 -0.008781646  0.3454251 -0.91480770 -0.161720638
## v203  0.010814648  0.5870280  0.36684838 -0.684778522
## plow -0.008377037 -0.1407382  0.06572953 -0.029551600
##           PC5          PC6
## v586  0.894858635  0.020170607
## v890 -0.013098760  0.007838988
## v204 -0.376805873  0.110982327
## v205  0.081416121  0.104637818
## v203  0.224488395  0.037239683
## plow  0.007082071  0.987359485
## Importance of components:
##           PC1          PC2          PC3          PC4          PC5          PC6
## Standard deviation    14.1100  1.47322  1.07065  0.61945  0.54252  0.31281
## Proportion of Variance  0.9799  0.01068  0.00564  0.00189  0.00145  0.00048
## Cumulative Proportion  0.9799  0.99054  0.99618  0.99807  0.99952  1.00000
```

Group 2: Violence

Violence

variable	description	N	type
v1649	Frequency of Internal Warfare (Resolved Rating)	152	ordinal
v1666	Individual Aggression - Assault	113	ordinal
v625	High Value Placed on Males Being Aggressive, Strong, and Sexually Potent	81	ordinal
mort	Male mortality during warfare (Ember)	80	continuous

Based on the VIFs, we excluded mort, male mortality during warfare.

```
##   v1649   v1666   v625   mort
## 2.991529 1.276236 1.064998 3.369642
```


The PCA on N=62 societies with complete cases indicated that the first component, consisting largely of v1649 (loading = 0.9917), expresses 79.7% of the variance, and the second, consisting largely of v1666 (loading = -0.9916) explains 19% of the variance across the sample, so we decided to include both these predictors, given that they are not collinear, and exclude v625 (loadings = |0.03|).

```
## Standard deviations (1, .., p=3):
## [1] 5.8631435 2.8648815 0.7297252
##
## Rotation (n x k) = (3 x 3):
##           PC1          PC2          PC3
## v1649  0.99170275  0.12499668 -0.03002465
## v1666  0.12383249 -0.99157785 -0.03793263
## v625   -0.03451323  0.03389987 -0.99882913

## Importance of components:
##           PC1      PC2      PC3
## Standard deviation    5.8631 2.8649 0.72973
## Proportion of Variance 0.7973 0.1904 0.01235
## Cumulative Proportion 0.7973 0.9877 1.00000
```

Group 3: Stratification

Social stratification

variable	description	N	type
v158	Scale 10- Social Stratification	186	ordinal
v2021	Distribution of Wealth	57	ordinal

The VIF analysis showed that the two measures, wealth inequality and social stratification are minimally collinear.

```
##      v158      v2021
## 1.048046 1.048046
```

Based on the PCA on N=58 societies, we decided to retain both variables, as they load largely independently on each PC (v2021 loading on PC1 = 0.9875, v158 loading on PC2 = -0.9875) and the first component explains 88.8% of the variance with the second explaining a non-negligible 11.1%.

```
## Standard deviations (1, .., p=2):
## [1] 0.8058769 0.2859464
##
## Rotation (n x k) = (2 x 2):
##           PC1          PC2
## v158  0.1575977 -0.9875034
## v2021 0.9875034  0.1575977

## Importance of components:
##           PC1      PC2
## Standard deviation    0.8059 0.28595
## Proportion of Variance 0.888 0.112
## Cumulative Proportion 0.888 1.000
```

```
## Standard deviation      0.8059 0.2859
## Proportion of Variance 0.8882 0.1118
## Cumulative Proportion  0.8882 1.0000
```

Group 4: Climate

Climate

variable	description	N	type
v200	Region	186	categorical
v857	Climate Type - Ordered in Terms of Open Access to Rich Ecological Resources D. White and M. Burton L986	186	ordinal
v186	Mean Annual Temperature (Xc)	180	continuous
v179	Latitude of Weather Station	179	ordinal
v180	Latitude Hemisphere (for Weather Stations)	180	binomial

The VIFs of v200 and v857 were both > 3, and still > 3 when removing one or the other of these two variables, based on collinearity with others. Therefore, we removed both of these variables.

```
##          GVIF Df GVIF^(1/(2*Df))
## v186  2.778129  1      1.666772
## v200  6.852734  5      1.212234
## v857  7.059428  5      1.215841
## v179  3.348937  1      1.830010
## v180  2.019411  1      1.421060

##          GVIF Df GVIF^(1/(2*Df))
## v186  2.352954  1      1.533934
## v857  3.688273  5      1.139416
## v179  2.816294  1      1.678182
## v180  1.133219  1      1.064527

##          GVIF Df GVIF^(1/(2*Df))
## v186  2.086532  1      1.444483
## v200  3.580283  5      1.136035
## v179  2.439087  1      1.561758
## v180  1.869974  1      1.367470
```

The PCA on N=179 societies indicated that we could use v186 (temperature) alone to express climate in our model, given that the first component explains 99.1% of the variance across populations, and is primarily v186 (loading = -0.997, other vars < |0.07|).

```
## Standard deviations (1, .., p=3):
## [1] 9.2751156 0.7648978 0.4358286
##
## Rotation (n x k) = (3 x 3):
##          PC1          PC2          PC3
## v186      -0.99734460  0.07282468  0.0005539704
```

```
## v179          0.07204268  0.98769110 -0.1388385290
## as.numeric(v180) -0.01065802 -0.13842995 -0.9903148772

## Importance of components:
##              PC1      PC2      PC3
## Standard deviation  9.2751 0.76490 0.43583
## Proportion of Variance 0.9911 0.00674 0.00219
## Cumulative Proportion 0.9911 0.99781 1.00000
```

Independent predictors

Independently treated predictors

variable	description	N	type
v53	Role of Father, Infancy	154	ordinal
v64	Population Density (from Murdock and Wilson Data)	184	ordinal
v714	Sex Ratio	90	ordinal
v740	Marriage Arrangements (Female)	151	ordinal
v1260	Pathogen stress: Total Pathogen Stress	186	ordinal

We included five additional variables in our model, not grouped with other predictors above. The VIF's on these five predictors were all <3.

```
##          v53          v64          v714          v740          v1260
## 1.049127  1.104785  1.063744  1.076269  1.072637
```

Note that in addition to v714, which is a three-level variable (female-biased, roughly even, male-biased), there also exists a continuous sex ratio variable (v1689) in the SCCS; however, we chose to include v714 because it has fewer missing values. Note though that sex ratio is stated to be subject to errors in the codebook.

Therefore, our final model consisted of v890 (female contribution to agriculture), v1649 (frequency of internal warfare), v1666 (assault frequency), v2021 (wealth inequalities), v158 (degree of social stratification), v186 (temperature), v53 (role of father during infancy), v64 (population density), v714 (sex ratio), v740 (arranged female marriages) and v1260 (pathogen stress).

Table S1: Summary of included variables

Variable	Original values (N)	Recoding and transformations
<i>Measures of polygyny</i>		
Cultural rules (v860)	1 = Monogamy prescribed (27), 2 = monogamy preferred (32), 3 = Polygyny preferred by leaders (45), 4 = Polygyny preferred by high social class (34), Polygyny preferred by most men (46), Missing data (2)	
% of married men polygynous (v871)	0 – 90, Missing data (41)	
<i>Predictors</i>		
Father role, infancy (v53)	1 = Distant (8), 2 = Rarely close (27), 3 = Occasionally close (72), 4 = Frequently close (44), 5 = Regularly close (3), Missing data (32)	Centered
Population density (v64)	1 = < 1 person per 5 sq. mile (36), 2 = 1 person per 1-5 sq. mile (22), 3 = 1-5 persons per sq. mile (25), 4 = 1-25 persons per sq. mile (27), 5 = 26-100 persons per sq. mile (34), 6 = 101-500 persons per sq. mile (21), 7 = over 500 persons per sq. mile (19), Missing data (2)	Converted to person per sq. mile, taking mean if range given – 1 = 0.2, 2 = 0.4, 3 = 3, 4 = 13, 5 = 63, 6 = 300.5, 7 = 500 – then centered
Stratification (v158)	1 = Egalitarian (65), 2 = Hereditary slavery (52), 3 = Two social classes, no castes/slavery (19), 4 = Two social classes, castes/slavery (20), 5 = Three social classes or castes, with or without slavery (30)	Condensed to three levels –1 + 2 = 1, 3 + 4 = 2, 5 = 3 – treated as factor
Temperature (v186)	23 (-16 – 29)	Standardized
Sex ratio (v714)	1 = Female excess (10), 2 = Roughly equal (60), 3 = Male excess (20), Missing data (96)	Centered
Arranged female marriages (v740)	1 = Individual selects partner autonomously, no approval required (12), 2 = Individual selects partner autonomously, approval required (40), 3 = Individual or others suggest match, then courtship (4), 4 = Individual choice and arranged marriages are alternatives (27), 5 = Parents choose, individual can object (35), 6 = Parents choose, individual cannot object (33), Missing data (35)	Recoded as binary – No = 1-5, Yes = 6
Female contribution to agriculture (v890)	0 = 0-4% (5), 1 = 5-14% (13), 2 = 15-24% (27), 3 = 25-34% (45), 4 = 35-44% (48), 5 = 45-54% (34), 6 = 55-64% (8), 7 = 65-74% (4), 8 = 75-84% (1), Missing data (1)	Converted to numerical scale, taking mean of ranges – 0 = 2, 1 = 9.5, 2 = 19.5, 3 = 29.5, 4 = 39.5, 5 = 49.5, 6 = 59.5, 7 = 69.5, 8 = 79.5 – then centered
Pathogen stress (v1260)	7– 21	Converted to natural logarithm, then standardized

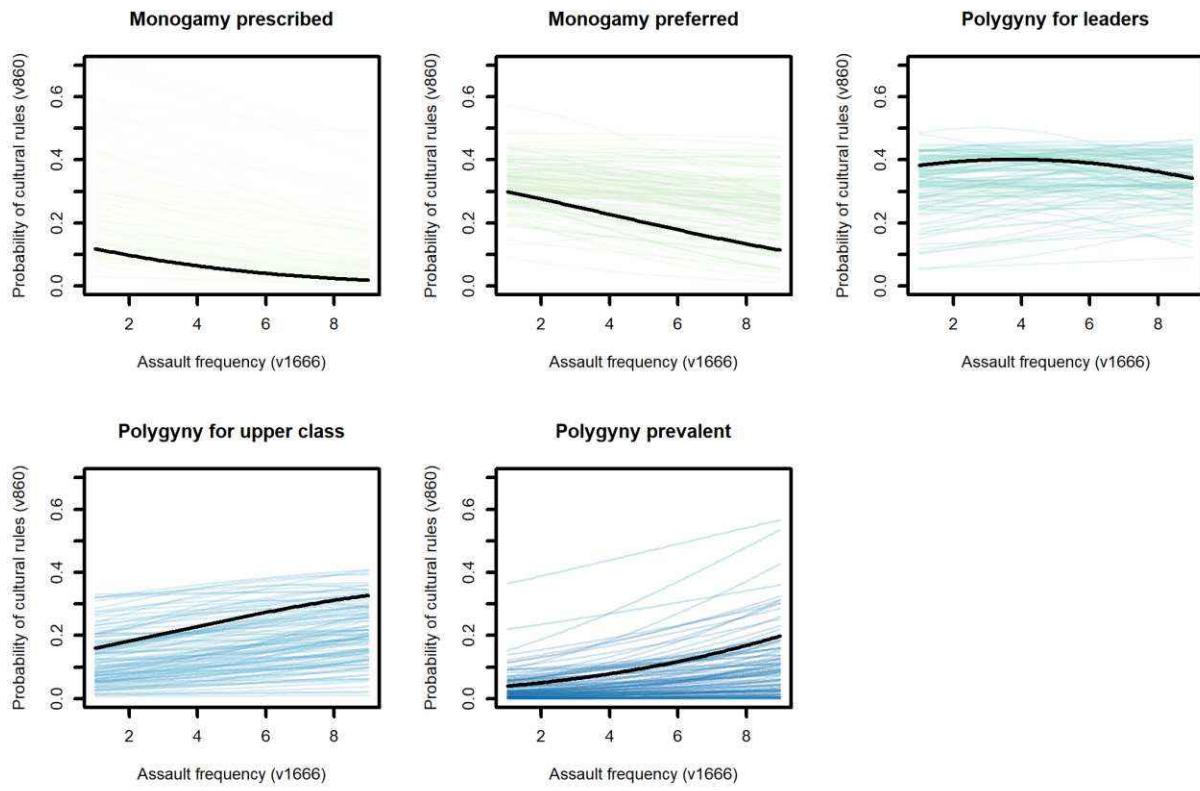
Internal warfare (v1649)	1 = Internal warfare absent or rare (60), 5 = Internal warfare occurs once every 3-10 years (7), 9 = Internal warfare occurs once every 2 years (4), 13 = Internal warfare occurs every year, seasonally (8), 17 = Internal warfare occurs almost constantly (27), Others in between these levels (46), Missing data (34)	Centered
Assault frequency (v1666)	1 = Low (29), 5 = Moderate (20), 9 = High (25), Others in between these levels (37), Missing data (73)	Centered
Wealth inequalities (v2021)	1 = General equality (17), 1.5 = Minor inequality (6), 2 = Some differences in wealth (17), 2.5 = Differences in wealth (3), 3 = Considerable differences in wealth (14), Missing data (129)	Condensed to three levels – 1 + 1.5 = 1, 2 + 2.5 = 2, 3 = 3 – treated as factor

Table S2: Summary of phylogenetic and non-phylogenetic models

Mean, lower (lwr) and upper (upr) 95% Highest Posterior Density Intervals, and posterior probability (PP) for all fixed effects parameters in the non-phylogenetic and phylogenetic models (see Figure 3 for full posterior probability distributions).

	Non-phylogenetic model				Phylogenetic model			
	Mean	Lwr 95%	Upr 95%	PP	Mean	Lwr 95%	Upr 95%	PP
<i>Outcome = v871</i>								
Intercept	1.39	0.81	1.98	1.00	1.17	-0.28	2.56	0.95
v1260	0.40	0.04	0.75	0.99	0.40	0.01	0.80	0.98
v740 = 2	0.54	-0.12	1.18	0.95	0.52	-0.13	1.15	0.94
v64	0.00	0.00	0.00	0.99	0.00	0.00	0.00	0.89
v53	-0.36	-0.68	-0.05	0.99	-0.22	-0.54	0.11	0.90
v186	0.17	-0.21	0.54	0.81	0.16	-0.24	0.57	0.78
v158 = 2	-0.36	-1.15	0.41	0.82	-0.23	-1.01	0.52	0.71
v158 = 3	-1.31	-2.35	-0.29	0.99	-1.08	-2.13	-0.05	0.98
v2021 = 2	0.34	-0.42	1.07	0.81	0.40	-0.37	1.14	0.85
v2021 = 3	0.82	-0.32	1.97	0.91	0.94	-0.16	2.04	0.95
v1649	-0.01	-0.06	0.03	0.71	-0.01	-0.05	0.04	0.61
v1666	0.12	0.01	0.22	0.98	0.10	0.00	0.20	0.97
v890	0.02	0.00	0.03	0.95	0.02	0.00	0.04	0.96
V714	0.43	-0.23	1.08	0.91	0.28	-0.38	0.96	0.80
Residual variance	1.73	1.22	2.27		1.42	0.96	1.92	
Phylo variance					1.82	0.27	4.13	
<i>Outcome = v860</i>								
Intercept	1.44	0.77	2.11	1.00	1.51	-0.32	3.31	0.95
v1260	0.50	0.11	0.90	0.99	0.50	0.00	0.99	0.98
v740 = 2	0.55	-0.17	1.28	0.93	0.60	-0.20	1.41	0.93
v64	0.00	0.00	0.00	0.99	0.00	0.00	0.00	0.92
v53	-0.34	-0.66	0.00	0.98	-0.17	-0.54	0.22	0.81
v186	0.24	-0.15	0.62	0.89	0.27	-0.20	0.74	0.87
v158 = 2	0.34	-0.53	1.20	0.79	0.49	-0.48	1.40	0.85
v158 = 3	-1.47	-2.61	-0.37	1.00	-1.37	-2.66	-0.14	0.99
v2021 = 2	0.69	-0.17	1.57	0.94	0.83	-0.17	1.80	0.95
v2021 = 3	1.29	-0.05	2.67	0.97	1.57	0.11	3.09	0.99
v1649	0.01	-0.04	0.06	0.64	0.02	-0.03	0.08	0.79
v1666	0.16	0.05	0.27	1.00	0.16	0.04	0.29	1.00
v890	0.01	-0.01	0.03	0.86	0.01	-0.01	0.03	0.83
v714	0.22	-0.56	0.94	0.74	0.02	-0.87	0.86	0.52
Cutpoint 1	1.21	0.85	1.57		1.38	0.95	1.82	
Cutpoint 2	2.52	2.07	3.00		2.86	2.29	3.45	
Cutpoint 3	3.62	3.09	4.18		4.15	3.46	4.88	
Residual variance	1 (fixed)				1 (fixed)			
Phylo variance					3.21	0.36	7.85	

Figure S1



Illustrating the predicted association between cultural rules constraining polygyny (v860) and assault frequency (v1666). Graphs plot the mean prediction (thick black line) and 100 samples randomly drawn from the posterior to illustrate the degree of uncertainty

Supertree inference and time-calibration

A matrix representation with parsimony (MRP) supertree of the Standard Cross-Cultural Sample (SCCS) populations is based on an updated and extended version of the dataset presented in Duda & Zrzavý (2016) and Duda & Zrzavý (in press). The dataset includes 388 source trees based on genetic or linguistic data from 251 studies published between 1990 and 2017 and includes 1,979 human populations.

The supertree was rooted with “all-0 outgroup” and two non-human outgroups (*Pan troglodytes*, *Homo neanderthalensis*). The dataset was analyzed using ‘new technology search’ in TNT version 1.5 (Goloboff & Catalano, 2016; Goloboff, Farris, & Nixon, 2008), using sectorial, ratchet, and tree fusing searches, obtaining trees from a 10,000 replicate random addition sequence, followed by branch swapping using the tree-bisection and reconnection method. The resulting supertree is a semi-strict consensus tree of 1,979 human populations, reduced to 186 populations present in the SCCS but retaining the original topology. Consequently, the relationships between the SCCS populations in the supertree are based on a much larger sample of populations for which more information is available. The analysis of the dataset that includes only the SCCS populations leads to a poorly resolved tree due to a high amount of missing data; 41 SCCS populations were entirely missing in the source trees and 33 populations were represented by the most suitable proxy population (i.e. neighboring or linguistically related population; Table S3). Eight populations (Kafa (Kaffa), Kimam, Yurok, Pomo (Eastern), Klamath, Kutenai, Natchez, Timbira) were placed solely based on linguistic classification (i.e. characters in the scaffold tree; see below).

The dataset contains high proportion of missing data due to presence of highly non-overlapping source trees (i.e. source trees with different taxon sets). This causes a loss in resolution of the resulting supertree (Bininda-Emonds & Sanderson, 2001). Moreover, the dataset is biased towards genetic information, as linguistic source trees are lacking for several language families. In order to mitigate these factors and to improve the resolution and topological stability of the supertree we used a ‘seed tree’ *sensu* (Bininda-Emonds & Sanderson, 2001), based on linguistic classification (Ruhlen, 1991; Simons & Fennig, 2017) (Table S4). The seed tree was upweighted by a factor of 1,000 to act as ‘scaffold tree’ *sensu* (Springer, Teeling, Madsen, Stanhope, & de Jong, 2001).

The scaffold tree allowed us to infer a stable phylogenetic placement of populations that are underrepresented or missing in the supertree dataset. It also ensures that linguistic information affects the tree globally. Indeed, there is a reason to believe that linguistic trees might be preferable to genetic trees as tools for phylogenetic comparative analyses of cross-cultural data, as cultural transmission may be largely shaped by the boundaries of language transmission. Languages track the inheritance of culture, and it is this inheritance that is under most circumstances pertinent to questions of human cultural evolution (Grollemund et al., 2015). In support of this, the resulting topology of the supertree is consistent with well-known expansions of language families associated with cultural traits, e.g. Bantu languages and farming traditions (Holden, 2002) or Austronesian seafaring traditions (Gray & Jordan, 2000).

Importantly, the scaffold tree is not an absolute constraint on monophyly of the language families, and it can be overruled by character information based on the source trees. Furthermore, the phylogenetic placement of populations speaking language isolates (e.g. Basques or Burusho) and populations which are single representatives of small language families (e.g. Hadza, Andamanese, or Haida), is based solely on their genes. The same applies to the populations of primary hunter-gatherers known to have recently adopted a language of unrelated neighboring populations (e.g. Mbuti Pygmies speaking a Nilo-Saharan language).

The SCCS supertree was time-calibrated by the `chronos` function in the *ape* package version 5.0 (Paradis et al., 2009) using autocorrelated relaxed-clock model (Paradis, 2013) and a set of non-overlapping time-constraints for 57 nodes and six tips (Table S5). The node age constraints correspond to divergence events inferred from genetic and linguistic studies of human population history; they include both deep and shallow nodes and are distributed evenly across the tree. The tip age constraints correspond to the focal years for the extinct populations in the SCCS (White, 2009).

References

- Arredi B., Poloni E.S., Paracchini S., Zerjal T., Fathallah D.M., Makrelouf M., . . . Tyler-Smith C., A predominantly neolithic origin for Y-chromosomal DNA variation in North Africa. *American Journal of Human Genetics* **75** (2), 2004, 338–345.
- Bergström A., Oppenheimer S.J., Mentzer A.J., Auckland K., Robson K., Attenborough R., . . . Tyler-Smith C., A Neolithic expansion, but strong genetic structure, in the independent history of New Guinea. *Science* **357** (6356), 2017, 1160–1163.
- Bininda-Emonds, O. R. P., & Sanderson, M. J. (2001). Assessment of the accuracy of matrix representation with parsimony analysis supertree construction. *Systematic Biology*, *50*(4), 565–579. <http://doi.org/10.1080/10635150120358>
- Callegari-Jacques S.M., Tarazona-Santos E.M., Gilman R.H., Herrera P., Cabrera L., dos Santos S.E.B., . . . Salzano F.M., Autosomal STRs in native South America—testing models of association with geography and language. *American Journal of Physical Anthropology* **145** (3), 2011, 371–381.
- Černý V., Fernandes V., Costa M.D., Hájek M., Mulligan C.J. and Pereira L., Migration of Chadic speaking pastoralists within Africa based on population structure of Chad Basin and phylogeography of mitochondrial L3f haplogroup. *Bmc Evolutionary Biology* **9**, 2009, 9.
- Chang W., Cathcart C., Hall D. and Garrett A., Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis. *Language* **91** (1), 2015, 194–244.
- Chaubey G., Metspalu M., Choi Y., Magi R., Romero I.G., Soares P., . . . Kivisild T., Population genetic structure in Indian Austroasiatic speakers: the role of landscape barriers and sex-specific admixture. *Molecular Biology and Evolution* **28** (2), 2011, 1013–1024.
- Chen Y.S., Olckers A., Schurr T.G., Kogelnik A.M., Huoponen K. and Wallace D.C., mtDNA variation in the South African Kung and Khwe—and their genetic relationships to other African populations. *American Journal of Human Genetics* **66** (4), 2002, 1362–1383.
- Ember, M. (1974). Warfare, sex ratio, and polygyny. *Ethnology*, *13*(2), 197–206.
- de Filippo C., Barbieri C., Whitten M., Mpoloka S.W., Gunnarsdottir E.D., Bostoen K., . . . Pakendorf B., Y-Chromosomal Variation in Sub-Saharan Africa: Insights Into the History of Niger-Congo Groups. *Molecular Biology and Evolution* **28** (3), 2011, 1255–1269.
- Fu Q. M., Mittnik A., Johnson P.L.F., Bos K., Lari M., Bollongino R., . . . Krause J., A Revised Timescale for Human Evolution Based on Ancient Mitochondrial Genomes. *Current Biology* **23** (7), 2013, 553–559.
- Goloboff, P. A., & Catalano, S. A. (2016). TNT version 1.5, including a full implementation of phylogenetic morphometrics. *Cladistics*, *32*(3), 221–238. <http://doi.org/10.1111/cla.12160>
- Goloboff, P. A., Farris, J. S., & Nixon, K. C. (2008). TNT, a free program for phylogenetic analysis. *Cladistics*, *24*(5), 774–786. <http://doi.org/10.1111/j.1096-0031.2008.00217.x>
- Gray, R. D., & Jordan, F. M. (2000). Language trees support the express-train sequence of Austronesian expansion. *Nature*, *405*(6790), 1052–1055. <http://doi.org/10.1038/35016575>
- Gray R.D., Drummond A.J. and Greenhill S.J., Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* **323** (5913), 2009, 479–483.
- Grollemund, R., Branford, S., Bostoen, K., Meade, A., Venditti, C., & Pagel, M. (2015). Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(43), 13296–301. <http://doi.org/10.1073/pnas.1503793112>
- Hammer M.F., Karafet T.M., Park H., Omoto K., Harihara S., Stoneking M. and Horai S., Dual origins of

- the Japanese: common ground for hunter-gatherer and farmer Y chromosomes. *Journal of Human Genetics* **51** (1), 2006, 47–58.
- Holden, C. J. (2002). Bantu language trees reflect the spread of farming across sub-Saharan Africa: A maximum-parsimony analysis. *Proceedings of the Royal Society B - Biological Sciences*, *269*(1493), 793–9. <http://doi.org/10.1098/rspb.2002.1955>
- Holman E.W., Brown C.H., Wichmann S., Muller A., Velupillai V., Hammarstrom H., . . . Egorov D., Automated Dating of the World's Language Families Based on Lexical Similarity. *Current Anthropology* **52** (6), 2011, 841–875.
- Honkola T., Vesakoski O., Korhonen K., Lehtinen J., Syrjanen K. and Wahlberg N., Cultural and climatic changes shape the evolutionary history of the Uralic languages. *Journal of Evolutionary Biology* **26** (6), 2013, 1244–1253.
- Hruschka D.J., Branford S., Smith E.D., Wilkins J., Meade A., Pagel M. and Bhattacharya T., Detecting Regular Sound Changes in Linguistics as Events of Concerted Evolution. *Current Biology* **25** (1), 2015, 1–9.
- Iriarte J., Smith R.J., Gregorio de Souza J., Mayle F.E., Whitney B.S., Cárdenas M.L., . . . Valdes P., Out of Amazonia: Late-Holocene climate change and the Tupi–Guarani trans-continental expansion. *The Holocene* **27** (7), 2016, 967–975.
- Kitchen A., Ehret C., Assefa S. and Mulligan C.J., Bayesian phylogenetic analysis of Semitic languages identifies an Early Bronze Age origin of Semitic in the Near East. *Proceedings of the Royal Society B: Biological Sciences* **276** (1668), 2009, 2703–2710.
- Li S., Schlebusch C. and Jakobsson M., Genetic variation reveals large-scale population expansion and migration during the expansion of Bantu-speaking peoples. *Proceedings of the Royal Society B-Biological Sciences* **281** (1793), 2014, 1148.
- Llamas B., Fehren-Schmitz L., Valverde G., Soubrier J., Mallick S., Rohland N., . . . Haak W., Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. *Science Advances* **2** (4), 2016, e1501385.
- Malaspinas A.-S., Westaway M.C., Muller C., Sousa, V.C., Lao O., Alves I., . . . Willerslev E., A genomic history of Aboriginal Australia. *Nature* **538** (7624), 2016, 207–214.
- Minocher, R. (2017). *Polygyny in the Standard Cross-Cultural Sample: A phylogenetically controlled re-analysis*. Ludwig-Maximilians University Munich.
- Olivieri A., Achilli A., Pala M., Battaglia V., Fornarino S., Al-Zahery N., . . . Torroni A., The mtDNA Legacy of the Levantine Early Upper Palaeolithic in Africa. *Science* **314** (5806), 2006, 1767–1770.
- Paradis, E. (2013). Molecular dating of phylogenies by likelihood methods: A comparison of models and a new information criterion. *Molecular Phylogenetics and Evolution*, *67*(2), 436–444. <http://doi.org/10.1016/j.ympev.2013.02.008>
- Paradis, E., Bolker, B. M., Claude, J., Cuong, H. S., Desper, R., Durand, B., ... de Vienne, D. (2009). Analyses of phylogenetics and evolution.
- Pickrell J.K., Patterson N., Barbieri C., Berthold F., Gerlach L., Guldemann T., . . . Pakendorf B., The genetic prehistory of southern Africa. *Nature Communications* **3**, (2012), 1143.
- Quinlan, R. J., & Quinlan, M. B. (2007). Evolutionary ecology of human pair-bonds: Cross-cultural tests of alternative hypotheses. *Cross-Cultural Research*, *41*, 149–169. <http://doi.org/10.1177/1069397106298893>
- Ramallo V., Bisso-Machado R., Bravi C., Coble M.D., Salzano F.M., Hunemeier T. and Bortolini M.C., Demographic Expansions in South America: Enlightening a Complex Scenario With Genetic and Linguistic Data. *American Journal of Physical Anthropology* **150** (3), 2013, 453–463.

- Ruhlen, M. (1991). *Guide to the world's languages: Classification* (Vol. 1). Palo Alto: Stanford University Press.
- Schlebusch C.M., Skoglund P., Sjodin P., Gattepaille L.M., Hernandez D., Jay F., . . . Jakobsson M. Genomic variation in seven Khoe-San groups reveals adaptation and complex African history. *Science* **338** (6105), 2012, 374–379.
- Shriner D., Tekola-Ayele F., Adeyemo A. and Rotimi C.N. Genome-wide genotype and sequence-based reconstruction of the 140,000 year history of modern human ancestry. *Scientific Reports* **4**, 2014, 6055.
- Simons, G. F., & Fennig, C. D. (2017). *Ethnologue: Languages of the world* (Twentieth).
- Springer, M. S., Teeling, E. C., Madsen, O., Stanhope, M. J., & de Jong, W. W. (2001). Integrated fossil and molecular data reconstruct bat echolocation. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(11), 6241–6. <http://doi.org/10.1073/pnas.111551998>
- Tobler R., Rohrlach A., Soubrier J., Bover P., Llamas B., Tuke J., . . . Cooper A., Aboriginal mitogenomes reveal 50,000 years of regionalism in Australia. *Nature* **544** (7649), 2017, 180.
- White, D. R. (2009). Pinpointing sheets for the Standard Cross-Cultural Sample: Complete edition. *World Cultures eJournal*, *17*(1).