

# Bayesian Phylogenetic Modeling of Cultural Evolution under the Influence of Selection

## The Origins and Maintenance of Female Genital Modification across Africa

Cody T. Ross · Pontus Strimling · Karen  
Eriksen · Patrik Lindenfors · Monique  
Borgerhoff Mulder

Received: date / Accepted: date

**Abstract** We present formal evolutionary models for the origins and persistence of the practice of Female Genital Modification (FGMo). We then test the implications of these models using normative cross-cultural data on FGMo in Africa and Bayesian phylogenetic methods that explicitly model adaptive evolution. Empirical evidence provides some support for the findings of our evolutionary models that the *de novo* origins of the FGMo practice should be associated with social stratification, and that social stratification should place selective pressures on the adoption of FGMo; these result, however, are tempered by the finding that FGMo has arisen in many cultures that have no social stratification, and that forces operating orthogonally to stratification appear to play a more important role in the cross-cultural distribution of FGMo. To explain these cases, one must consider cultural evolutionary explanations in conjunction with behavioral ecological ones.

Cody T. Ross  
Department of Anthropology  
One Shields Avenue, University of California, Davis  
Davis, CA 95616 USA  
E-mail: ctross@ucdavis.edu

Pontus Strimling  
Centre for Studies of Cultural Evolution  
Stockholm University  
Stockholm, Sweden

Karen Eriksen  
Department of Psychology  
One Shields Avenue, University of California, Davis  
Davis, CA 95616 USA

Patrik Lindenfors  
Centre for Studies of Cultural Evolution  
Stockholm University  
Stockholm, Sweden

Monique Borgerhoff Mulder  
Department of Anthropology  
One Shields Avenue, University of California, Davis  
Davis, CA 95616 USA

We conclude with a discussion of the implications of our study for policies designed to end the practice of FGMo.

**Keywords** Female Circumcision · FGM · Cultural Evolution · Social Transmission · Marriage Markets · Phylogenetic Modeling

## 1 Introduction

In many parts of the world the prevailing culture requires that women undergo genital cutting/modification prior to their marriage. These operations typically occur between infancy and late puberty, and may be extremely costly in terms of health, survival, and reproduction; these operations thus constitute a major challenge for evolutionary explanation. The analyses presented herein test competing hypotheses for the distribution of this cultural trait, with a particular focus on how the conditions that may have selected for the emergence of the trait (arguments for origins) might differ from those that account for its persistence (arguments for maintenance). We first present a model that demonstrates how conditions selecting for the origins of female genital modification<sup>1</sup> (FGMo) might differ from those that select for its persistence over time. In particular, we model how wealth and/or status differences within social groups might select for the emergence of this potentially costly trait, and how subsequent frequency-dependent forces might keep the trait in the population, despite its costs. We then use empirical data from Africa to test whether status differences and social stratification are associated with the origins of FGMo. Our expectation (derived from the model) is that stratification will be associated with the *de novo* origins of FGMo, and will increase the likelihood of the inter-cultural transmission of FGMo. We use our results to discuss the adaptive significance of apparently costly cultural traits, and to discuss the importance of differentiating explanations for the origins and the maintenance of cultural traits.

Our study lies at the intersection of several literatures. First, there is a large literature on how marriage markets affect a range of cultural practices, including marriage payments [Anderson, 2003; Bell and Song, 1994; Borgerhoff Mulder, 1995] and other household outcomes [Chiappori et al, 2002; Quisumbing et al, 2000]. Many of these insights are inspired by Becker [1981]. Second, we develop hypotheses on the basis of a widely-ranging research literature on female “circumcision” [Ericksen, 1989; Shell-Duncan and Hernlund, 2000; Silverman, 2004]. While contributors to this literature span many different disciplines, there is considerable agreement over the hypothesized functions of FGMo and related practices, though

<sup>1</sup> Also known as: Female Circumcision, Female Genital Mutilation (FGM), Female Genital Cutting (FGC), or a combination of these terms like FGM/C. There has been concern over the terminology one should use when discussing this practice. We purposefully avoided using the term “mutilation” in the text of this paper, as we feel that the use of the term “mutilation” is unduly value-laden, and insensitive to the women who have endured the practice. Portraying females as mutilated beings is likely to be offensive and disrespectful to many, and may decrease the willingness of respondents to discuss the practice with researchers. Likewise, we feel that it is wrong to distance the practice of female genital modification from male genital modification (circumcision), as such an action seem to validate one type non-necessary, non-consensual, amputation of genital tissue (common in “Western” culture), while stigmatizing a similar practice in other cultures. We use the more neutral term FGMo to contextualize the practice within the wider anthropological scope of body modification.

less agreement with regards either to its specific geographical origins or to its meaning. Third, we draw indirectly on the insights of signaling (or handicap) theory, as developed in evolutionary biology and applied in anthropology [Grafen, 1990; Smith et al, 2003], viewing FGMO as a signal demonstrating sexual fidelity and paternity certainty. Fourth, we use formalized evolutionary models to identify distinct hypotheses for the origins and maintenance of costly cultural traits, as these potentially differing dynamics are rarely separated in analysis [Borgerhoff Mulder et al, 2006]. Finally, there is a growing (albeit widely-dispersed) literature across academic venues and policy circles that addresses the morality of, and challenges to, eradicating the practice of FGMO [Shell-Duncan and Hernlund, 2000]. We use our results to evaluate some of the debates in this complex area. Additionally, we argue that the statistical methods used in this case study have broad implications for advancing cultural phylogenetic analysis.

## 1.1 Female Genital Modification

The prevalence of FGMO worldwide is unknown, but it is estimated that more than 125 million girls and women alive today have undergone some form of FGMO, with another 30 million at risk of being cut in the next decade [UNFPA-UNICEF, 2013]. It is most common in Africa, affecting up to 90 percent of women in Djibouti, Egypt, Eritrea, Mali, Sierra Leone and Somalia, over 50 percent of women in other African countries such as Benin, Ethiopia, Burkina Faso, Chad, Ethiopia, Gambia, Guinea Bissau, Kenya, Liberia, Nigeria, Sudan, and Togo. It occurs elsewhere across southern and central Africa and in the Middle East, for example in Oman, Yemen, and the United Arab Emirates, as well as in some Asian countries including Indonesia, Malaysia, Sri Lanka and India [Toubia and Izett, 1998].

Little is known about the origins of the practice. Scholars have proposed a single origin in Egypt, on the the basis of circumcised 5th century BC mummies [Little, 2003], or ancient Meroe [Mackie, 2000; Mackie and LeJeune, 2009]. It is suggested that in these highly stratified ancient empires infibulation was practiced in the context of extreme resource inequality, with highly polygynous wealthy males cutting girls and women to signal the fidelity of their wives. Mackie has hypothesized that the practice diffused across social strata and spread along female slave trade routes. Others suggest a dual-source origin, arguing that as FGMO spread out of its original core areas, it encountered and merged with pre-existing practices associated with initiation rites for both males and females [Dorkenoo, 1994]. Strong evidence for these claims is nevertheless lacking. While the practice sometimes spread with Islam, there are many non-Islamic groups who practice FGMO.

Types of operations vary in severity, entailing either the removal of the prepuce or hood of the clitoris (clitoridectomy), removal of the clitoris and all or part of the labia minora (excision), or removal of the complete clitoris, labia minora and all or part of the labia majora followed by a sewing together of the labia (infibulation, or Pharonic circumcision).

The health costs (and benefits) of FGMO are hotly disputed [e.g. Shell-Duncan and Hernlund, 2000]. Estimates are likely to be biased by concealed complications when the practice is illegal (downwards), by prejudice and by proponents of elimination strategies (upwards), and most seriously by lack of good data. Fur-

thermore western observers tend to stress pain, reduced sexual enjoyment and medical complications, while (some) Africans emphasize the cultural importance of the tradition; for example a Kenyan woman commented that FGMo might be seen as “buying maturity with pain” [Davison, 1996, p60]. Nevertheless, there is clear evidence that these operations, and particularly the more severe forms, can cause extensive short- and long-term medical complications, with implications for maternal health, pregnancy loss and stillbirth, primary and secondary fertility, and child survival [Mackie, 2003; Shell-Duncan and Hernlund, 2000; Banks et al, 2006], as well as for mental health [e.g. Whitehorn et al, 2002].

Most pertinent to this paper are arguments regarding the function of FGMo, or similar practices with likely negative effects for women’s health and well being. There are three types of explanations. The first, and by far the most predominant, concerns the marriageability of women. Though specifics differ markedly in different parts of the world, virginity, ‘purity,’ and sexual restraint before marriage are highly regarded in many societies that practice FGMo. FGMo operations are seen as a way of protecting sexual propriety, morality, paternity, and marriageability [Erickson, 1989; Little, 2003; UNFPA-UNICEF, 2013]. Furthermore, FGMo is often associated with veiling, child betrothal, virginity testing at marriage, a transfer of sexual and reproductive rights to the groom and his family at marriage, and the legitimacy of subsequent births. There are cases where FGMo is found in groups that lack strong chastity and fidelity expectations, for example the Rendille women of Kenya are free to engage in premarital sex, but must undergo FGMo at marriage [Mackie, 2000]. Nevertheless, scholars investigating FGMo across multiple sites typically stress the theme of safeguarding female sexual purity, enhancing marriage chances, and preserving family honor [Shell-Duncan and Hernlund, 2000; UNFPA-UNICEF, 2013]. For instance, uninfibulated women in Sudan during the early 1970s were considered to be like prostitutes [Hayes, 1975].

A second set of explanations are more heterogeneous. They include protecting the health of a baby, increasing the likelihood of conception, reducing (or increasing) sexual pleasure, achieving an aesthetic ideal, or becoming a fully adult member of society; these more proximate considerations are also linked to marriageability, albeit indirectly. Particularly idiosyncratic notions include the belief that FGMo prevents child mortality (through avoiding fatal connection between baby’s head and the clitoris during delivery [Myers et al, 1985]), that FGMo cured certain ‘female psychological disorders’ in Victorian England [Little, 2003], and that genital cutting is a functionless fad, as witnessed recently in Chad. While FGMo in southern Chad may have been adopted in the 19th century to deter slave raiders from taking women, adolescent girls in the 1990s reportedly sought out the operation, often without parental consent, as something “fun, rebellious and cool” [Leonard, 2000, p190].

The third explanation stresses the importance of tradition, custom, and cultural identity. For instance, for the Kipsigis of Kenya their very name implies “we the circumcised”, referring to the rebirth that is believed to occur at circumcision ceremonies, a central part Kipsigis identity vis-a-vis other ethnic groups [Daniels, 1970]. Even more strategically, the Kenyan Mukogodo appear to have adopted circumcision ceremonies to hasten their transition to becoming Maasai [Cronk, 2004]. In a compilation of studies on attitudes toward FGMo more than half the respondents reported that tradition was their primary reason for undergoing the operation [Warzazi, 1986]. Some authors emphasize only the importance of follow-

ing the ancestors, others note strong social pressure to conform to the behavior of others, and yet others explicitly recognize the force of current social convention.

As many researchers on FGMO recognize, there are intricate interdependencies between these hypotheses. Indeed even in cases where marriageability is not the explicitly-stated motivation for FGMO, as in the Senegambian region, [Shell-Duncan and Hernlund, 2007]. It is extremely difficult to distinguish the more general motivations of social respectability and conformity from those of being suitable as a wife [Mackie and LeJeune, 2009]. Furthermore, some of the more idiosyncratic beliefs underlying the need for FGMO may also serve as a way of encouraging conformity to social norms. Accordingly Mackie [2000], emphasizing the critical role of social convention in maintaining particular practices, notes how the belief that an uncut woman will be unfaithful fixes the custom, even if the conditions that first gave rise to the trait have changed: “As soon as women believed that men would not marry an unmutated woman, and men believed that an unmutated would not be a faithful partner in marriage, the convention was locked in place” [Mackie, 2000, p264].

## 1.2 An Evolutionary Approach

Our evolutionary analysis of FGMO draws closely on the literature reviewed above. Parents or other kin may choose to subject young female relatives to costly operations to enhance marriageability. In other words, the trait may be sexually selected, functioning to enhance the access of females to favored mates. The intuition here is that in an initial non-FGMO population parents use FGMO to ensure and signal the virginity of their daughters at marriage. Insofar as women who have undergone FGMO observe more sexual restraint (either as a consequence of the operation or of associated customs), FGMO also signals higher paternity certainty to a prospective groom. In conditions where there is little difference in quality among prospective grooms it is unlikely that parents would inflict this physiological cost on their daughters. However where there is extensive competition among women to become the brides of particularly wealthy or powerful men, a costly trait like FGMO could arise as a bargaining tool. The simple logic here parallels evolutionary arguments for the origin of another costly display—dowry, and its association with stratification [Dickemann, 1979; Gaulin and Boster, 1990, *see also Fortunato et al.*, 2006] and has in fact been marshaled before as an explanation for the origins of FGMO [Mackie, 2000]. Some support comes from evidence that female puberty rituals are more elaborate and costly as social complexity increases, although this is based on a nonrandom sample drawn from a selection of foraging societies deemed to be weakly stratified [Hayden et al, 1997]. Other support comes from comparative analysis that links public menarcheal ceremonies to strong fraternal interest groups, where it appears that powerful kin-based units are making explicit contractual negotiations over women [Ericksen, 1989; Paige and Paige, 1981].

As noted above, other functions can accrete to the practice of FGMO over time. For example, FGMO (at least when accompanied by ritual) can constitute a key rite of passage [Van Gennep, 1960]; it is also a cardinal symbol of ethnicity, as groups that do not practice FGMO are sometimes seen as barely human by FGMO practicing groups. There are also clear reasons why it is difficult to dislodge FGMO

once it is in place as a marker of marriageability, since a family that abhors FGMO for their daughters will most likely fail to find husbands, and not contribute to the next genetic (or cultural) generation. There is a strong coordination system in place—without a critical mass of FGMO-rejecters, there is a strong disincentive for families to reject the practice.

This argument leads to the hypothesis that the *origins* of FGMO are likely to be linked to stratification, based on the rationale that where there exists considerable variance among men in mate quality, and where men invest heavily in their offspring, parents will inflict physiological (and other) costs on their daughters to reap the benefits of a marriage to a high-ranking man. Conversely, the custom might *persist* purely as a consequence of the high cost of not conforming, irrespective of the presence of high variance in wealth or status among men.

## 2 Modeling the Emergence and Persistence of FGMO

We begin methodologically by providing formal theoretical models for the emergence and persistence of FGMO. We then use phylogenetic models of adaptive evolution to test the implications of the theoretical models using cross-cultural data.

### 2.0.1 Assumptions and Notations

The following assumptions and notations are constant across all models:

a. Mating assessment is universal. That is—every woman agrees on a mating value for each man, and every man agrees on a mating value for each woman. A given woman's mating value is denoted  $f$  (and a given man's  $m$ ). The distributions for the mating values of women and men are  $d_f$  and  $d_m$ , respectively. Mating values are reflective of the integrated suite of traits of interest to the opposite sex.

b. The symbols  $r_f$  and  $r_m$  denote the functions that map the mating value of an individual woman or man to her or his rank in the mating market. These functions lead to a ranking of all women where the woman with the highest mating value gets rank  $r_f = 1$ , the woman with second highest mating value gets rank  $r_f = 2$ , and so on. The woman with rank  $r_f = 1$  is then assumed to marry the man with rank  $r_m = 1$  and so on. Thus, if  $f_1$  is the woman with the highest mating value in the  $d_f$  distribution, then her rank is described as  $r_f(f_1, d_f) = 1$ ; she will be paired with the male where  $r_m = 1$ , the highest ranked male from the  $d_m$  distribution.

c. The marriage value of a given person in  $d_f$  or  $d_m$  can be returned as a function,  $V_f$  or  $V_m$ , of rank, such that  $V_m(r_m(m_1, d_m))$ , for example, is the marriage value of the highest ranked man. The marriage value function is order preserving, but allows for a non-linear relationship between rank order and marriage value; a unit change in rank may not significantly impact marriage value near the bottom of a ranking scale, but might have a huge impact on marriage value near the top of a ranking scale—especially in stratified societies.

d. Everyone gets married.

e. FGMO comes at a constant cost,  $c$ .

f. FGMO gives a woman a higher mating value, and therefore, an opportunity to marry a more valuable man.

g. The sex ratio is balanced such that the number of women in the population,  $N_f$ , is equal the number of men in the population,  $N_m$ .

h. There is a benefit,  $s$ , to having the same behavior as the rest of the group. Generally,  $s$  is a function of the proportion of women,  $\alpha$ , that have the FGMO behavior. We conceive of this pressure as resulting from frequency-dependent dynamics, such as conformist or prestige bias [Henrich and Boyd, 1998; Henrich and Gil-White, 2001; McElreath et al, 2003, and see Discussion].

i. Baseline fitness across women is described as  $\bar{w}$ .

### 2.0.2 The General Model

We will consider two models. In the first, FGMO gives a virginity assurance, and therefore gives a constant increase in mating value,  $i$ , to every woman who has FGMO. In the second model, FGMO acts as a costly signal of general fitness. In this model, we see the increase in mating value as a random variable,  $x$ , since different women can signal differential fitness in a non-deterministic way. Figure 1 illustrates how the mathematical mechanisms outlined here function to produce the behavior of the model.

In both models a woman who does not undergo FGMO has payoff,  $P$ :

$$P = \bar{w} + V_m(r_f(f, d_f)) + s(1 - \alpha) \quad (1)$$

where  $\alpha$  is the proportion of the population who undergo FGMO. Equation 1 says that the value she gets from marrying is dependent on the man she marries, which is dependent on her rank  $r_f(f, d_f)$ , which is dependent on her value  $f$  and the distribution of mate values across all females  $d_f$ . The last term is the social payoff from having the same behavior as proportion  $1 - \alpha$  of the population.

### 2.0.3 The Virginity Assurance Model

In the virginity assurance model, the payoff,  $P_{VA}$ , for undergoing FGMO is:

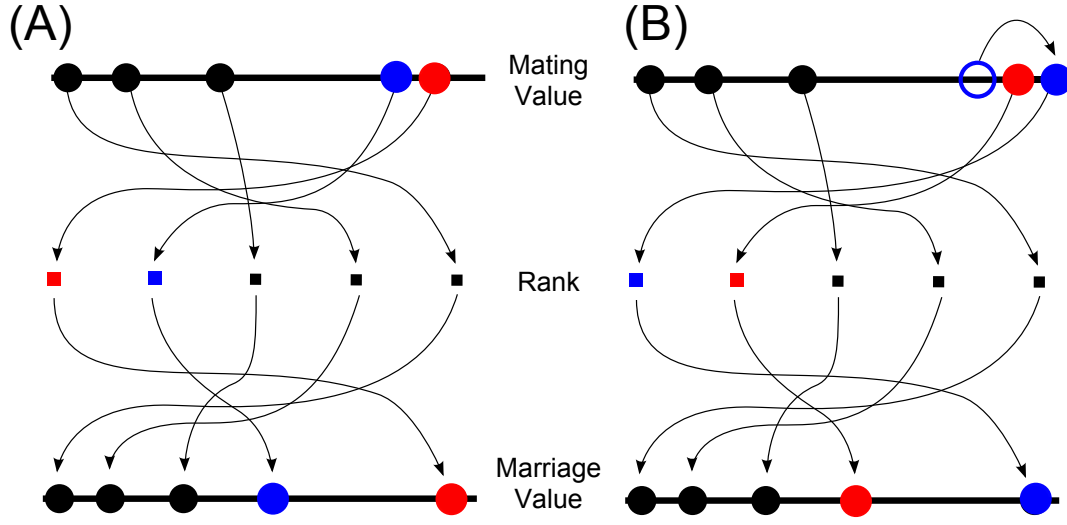
$$P_{VA} = \bar{w} + V_m(r_f(f + i, d_f)) + s(\alpha) - c \quad (2)$$

This gives us a condition that FGMO will emerge when there is a woman such that:  $P_{VA} - P > 0$ , for  $\alpha = 0$ . That is when:

$$V_m(r_f(f + i, d_f)) - V_m(r_f(f, d_f)) - s(1) - c > 0 \quad (3)$$

So for FGMO to emerge we need at least one (or a combination) of the following: low social control or conformity bias,  $s$ , a high value of FGMO,  $i$ , or a high difference in the marriage value of men,  $V_m$ . We will not elaborate on how FGMO spreads through the population, nor will we discuss all possible equilibria that can appear. Rather we will focus on what is needed for an equilibrium where everyone in the population does FGMO. In the virginity assurance model, this occurs for cases such that when  $\alpha = 1$ , then  $P_{VA} - P > 0$  holds for all women in the population. That is when:

$$V_m(r_f(f + i, d_f)) - V_m(r_f(f, d_f)) + s(1) - c > 0 \quad (4)$$



**Fig. 1** Figure 1 illustrates a simplified sketch of the mathematical model described in this section. In frame (A), we observe a distribution of mating values (on the real number line) for five women. We have highlighted two individuals in red (highest ranked) and blue (second highest ranked). The arrows from Mating Value to Rank are indicative of the rank function  $r_f(f, d_f)$  and map mating values (on the real number line), to an ordered ranking (on the integers). The arrows from Rank to Marriage Value are indicative of the value function  $V_m()$  and map a woman's rank score on the marriage market to the marriage value of her husband (on the real number line). In frame (B), we imagine that the woman shown in blue has undergone FGMo and has thus added  $i$  or  $x$  to her mating value. This allows her to move from her previous position on the Mating Value scale (shown in the blue outlined circle) passed the woman labeled in red. This in turn increases her rank and allows her access to the marriage value of the highest ranked man.

Here two things have changed from the emergence condition in Equation 3; the social effect has switched so that it now benefits FGMo, and also the value distribution of women  $d_f$  has changed since every woman has added  $i$  to her value. This means that when  $\alpha = 1$  every woman is in the same position on the ranking scale as she would have been had no one undergone FGMo.

So Equation 4 can be reduced to:

$$s(1) - c > 0 \quad (5)$$

There are several interesting conclusions to draw from Equation 5. Whereas the marriage value function for males,  $V_m$ , is critical in explaining the origins of FGMo, it has no effect on the *maintenance* of the trait at fixation. Another conclusion is that strong social pressure is required to obtain an equilibrium where every woman has undergone FGMo. Finally, we can see that the difference in payoff between being a woman in a society where everyone has FGMo and a society where no-one has FGMo is strictly negative for every woman, since the rankings are unchanged between the FGMo and non-FGMo equilibria, but all women pay a cost  $c$  in the FGMo equilibrium. Every man on the other hand increases his payoff by  $i$  in the FGMo equilibrium.



#### 2.0.4 The Costly Signaling Model

Here we assume that FGMO functions as a general costly signal of fitness, rather than a signal of virginity *per se*. It is unreasonable to imagine that this effect is the same for each woman or child, or that the effect is known before the decision to commit FGMO is taken. Therefore we model the mating value increase from having FGMO as a random variable  $x$ . The payoff,  $P_{CS}$ , to undergoing FGMO is thus:

$$P_{CS} = \bar{w} + V_m(r_f(f + x, d_f)) + s(\alpha) - c \quad (6)$$

And FGMO emerges in a population if the expected value of  $P_{CS}$  exceeds  $P$  for at least one woman in the population when  $\alpha = 0$ . That is, when:

$$E[V_m(r_f(f + x, d_f))] - V_m(r_f(f, d_f)) - s(1) - c > 0 \quad (7)$$

An equilibrium where everyone in the population undergoes FGMO is maintained if  $E[P_{CS}] > P$  for every woman in the population when  $\alpha = 1$ . That is when:

$$E[V_m(r_f(f + x, d_f))] - V_m(r_f(f, d_f)) + s(1) - c > 0 \quad (8)$$

And here the difference from the virginity assurance model becomes clear. If we again look at a case in which a low ranking woman in a non-FGMO society considers FGMO, she can (depending on the distributions of  $x$ ,  $d_f$  and  $d_m$ ) still have:

$$E[V_m(r_f(f + x, d_f))] - V_m(r_f(f, d_f)) > 0 \quad (9)$$

Even though the probability of having a higher payoff after FGMO might be small, FGMO could have a high enough expected value to be worthwhile. This also means that there might be women who benefit from living in a society where everyone undergoes FGMO, which was not the case in the virginity assurance model. On average though, each woman still gets  $c$  less payoff, while each man increases his by an average value of  $x$ .

For a better understanding of our results let's look closer at the implications of Equation 9. For the woman with lowest mating value before FGMO, Equation 9 can be expressed as:

$$\sum_{k=0}^N \Phi(k) V_m(N_f - k) - V_m(N_f) > 0 \quad (10)$$

where  $\Phi(k)$  is the chance that undergoing FGMO gives her a higher value than  $k$  other women, and  $N_f$  is the number of women in the population. Note that as long as  $\Phi(k)$  is not strictly zero for all  $k$ , then the left-hand side of Equation 10 is strictly positive, although it is neither necessarily greater than  $c$ , nor necessarily large enough to move a woman up on the mating value ranking scale. Thus, FGMO could function to increase a woman's mating value relative to other women, though this is not guaranteed generally.

To further investigate the implications of this model, we make some reducing assumptions that  $d_f$  and  $x$  are normally distributed random variables. In this case, all that matters to our models are the variances of the respective distributions; to understand the impact of variance in  $x$  and  $d_f$ , let us note that the probability that a realization from  $\text{Normal}(\mu_1, \sigma^2)$  is greater than a realization from  $\text{Normal}(\mu_2, \sigma^2)$

is just the probability that  $\text{Normal}(\mu_1 - \mu_2, 2\sigma^2) > 0$ , which is determined by the difference between  $\mu_1$  and  $\mu_2$  and by  $\sigma$ . In our model, the difference between  $\mu_1$  and  $\mu_2$  is determined by the variance in the mating value of women,  $d_f$ , and  $\sigma$  is the variance of the random variable  $x$ .

An increasing variance of  $d_f$  decreases the chance that a low ranking woman can move up on the rank scale as a function of FGMO adoption. An increase in the variance of  $x$  increases the chances that a low ranking woman can move up on the mating rank scale. Finally, as  $V_m$  becomes an accelerating function of increasing rank, it increases the value for a woman of moving up on the mating rank scale.

## 2.1 Conclusions of the Model

In conclusion, our formal evolutionary models show that variance in male mating value should be strongly associated with the origins of FGMO. Further, we show that variance in male mating value is not strictly required for the maintenance of FGMO at fixation, so long as social pressure is intense enough to maintain FGMO after it originates. Variance in male mating value may, however, factor into the maintenance of FGMO when the prevalence of FGMO is below fixation.

The costly signaling model suggests that some women can benefit from FGMO at fixation, whereas under the virginity assurance model no women benefit from FGMO at fixation. Because the costly signaling model requires data on variability in female quality, we restrict our empirical testing here to the virginity assurance model.

## 3 Methods

### 3.1 Adaptive Phylogenetic Analysis of Empirical Data

#### 3.1.1 Data Sources and Sample Selection

To test whether variance in male value is associated with the distribution of FGMO across cultures, we use an extensive coding of data on female genital modification in Africa [Erickson, 1989], the continent in which the custom is largely concentrated.

The sample consists of 113 cultural clusters in Africa, drawn from Murdock's Ethnographic Atlas [Murdock, 1969]. For this sample, Erickson [1989] examined all available ethnographic sources for evidence of FGMO. Following Erickson [1989], clusters that contained more than one culture (see Murdock [1968] for further discussion) are represented by the culture with the best coverage on the topics being coded; in cases where more than one culture had good source materials, the choice was determined randomly.

Female genital modification, *FGMO*, was coded as present if the practice was mentioned as occurring by at least one ethnographer, either present or past, irrespective of how much detail was given. We use the existence of economic and/or social stratification as our measure of variation in male quality, on the assumption that in more highly stratified societies there are greater differences in male resource holdings, and greater power and prestige differences; stratification is coded using Variable 67 (class stratification) in Murdock [1957]. For a small number of

cultural groups, the Ethnographic Atlas had missing information; we investigated the primary ethnographic literature to obtain equivalent data for these groups. The Supplementary Materials (*FGMoData.csv*) contains our data and the relevant citations to primary sources.

### 3.1.2 Full Bayesian Phylogenetic Modeling of Adaptive Evolution

To model the effect of stratification on the evolution of FGMo, we utilize a full Bayesian phylogenetic model of adaptive evolution (based on an Ornstein-Uhlenbeck process) in the spirit of [Butler and King \[2004\]](#). The methods advanced by [Butler and King \[2004\]](#), conceptualize evolution across a phylogeny as a function of both selective processes and drift. This approach thus constitutes critical progress in phylogenetic analysis insofar as it allows for adaptive hypotheses to be evaluated with phylogenetic models that include selection dynamics explicitly. The methods advanced by [Butler and King \[2004\]](#) allow us to: 1) investigate the extent to which the likelihood of a cultural group practicing FGMo is conditioned on the state of that cultural group as stratified versus non-stratified, and 2) contrast the strength of selection for FGMo based on stratification, with the strength of drift (and selective forces operating orthogonally to stratification).

Standard tools for phylogenetic inference concerning the evolution of discrete traits [[Pagel, 1994](#); [Pagel and Meade, 2006](#)] represent an alternative analytic strategy. We focus our analysis in the main text on the the adaptive phylogenetic analysis, as we find the model dynamics to more elegantly match the empirical processes we wish to understand. For thoroughness, we present the results from a discrete trait phylogenetic model in the Supplementary Materials, Appendix A Section 1.7.

[Butler and King \[2004\]](#) developed a software package (OUCH, Ornstein-Uhlenbeck for Comparative Hypotheses) for adaptive phylogenetic analysis in the R software environment [[R Core Team, 2013](#)]. This software package, however, treats phylogenies as known data, deals poorly with parameter constraints, and relies on a fragile maximum likelihood estimation procedure, which fails to function reliably in many contexts, as detailed in [Butler and King \[2004\]](#).

To improve on the software introduced by [Butler and King \[2004\]](#), we wrote our own full Bayesian implementation of the adaptive phylogenetic model using Hamilton Markov Chain Monte Carlo (HMC) simulation [[Hoffman and Gelman, 2013](#)] in C++, using the Stan 2.2.0 library [[Stan Development Team, 2013a](#)]. Our Bayesian implementation allows us to integrate over phylogenetic uncertainty, impose parameter constraints, and use prior information to identify parameters that are not necessarily identifiable under maximum likelihood estimation.

Following [Butler and King \[2004\]](#), we imagine trait evolution over a phylogenetic tree occurring as an Ornstein-Uhlenbeck process where a trait,  $X$ , evolves under a regime composed of both selection and drift such that:

$$\partial X(t) = \alpha[\beta(t) - X(t)]\partial(t) + \sigma\partial B(t) \quad (11)$$

where  $\partial X(t)$  is the change in the character trait,  $X$ , over the course of a small increment of time,  $\alpha$  is the strength of selection,  $\beta(t)$  is the optimal trait value, and  $\sigma$  mediates the intensity of “white noise” fluctuations,  $\partial B(t)$ .

Below we describe the phylogeny utilized in our analysis and then outline the mathematical details of our adaptive phylogenetic model.

### 3.1.3 A Phylogeny of Language Families Based on Lexical Similarity

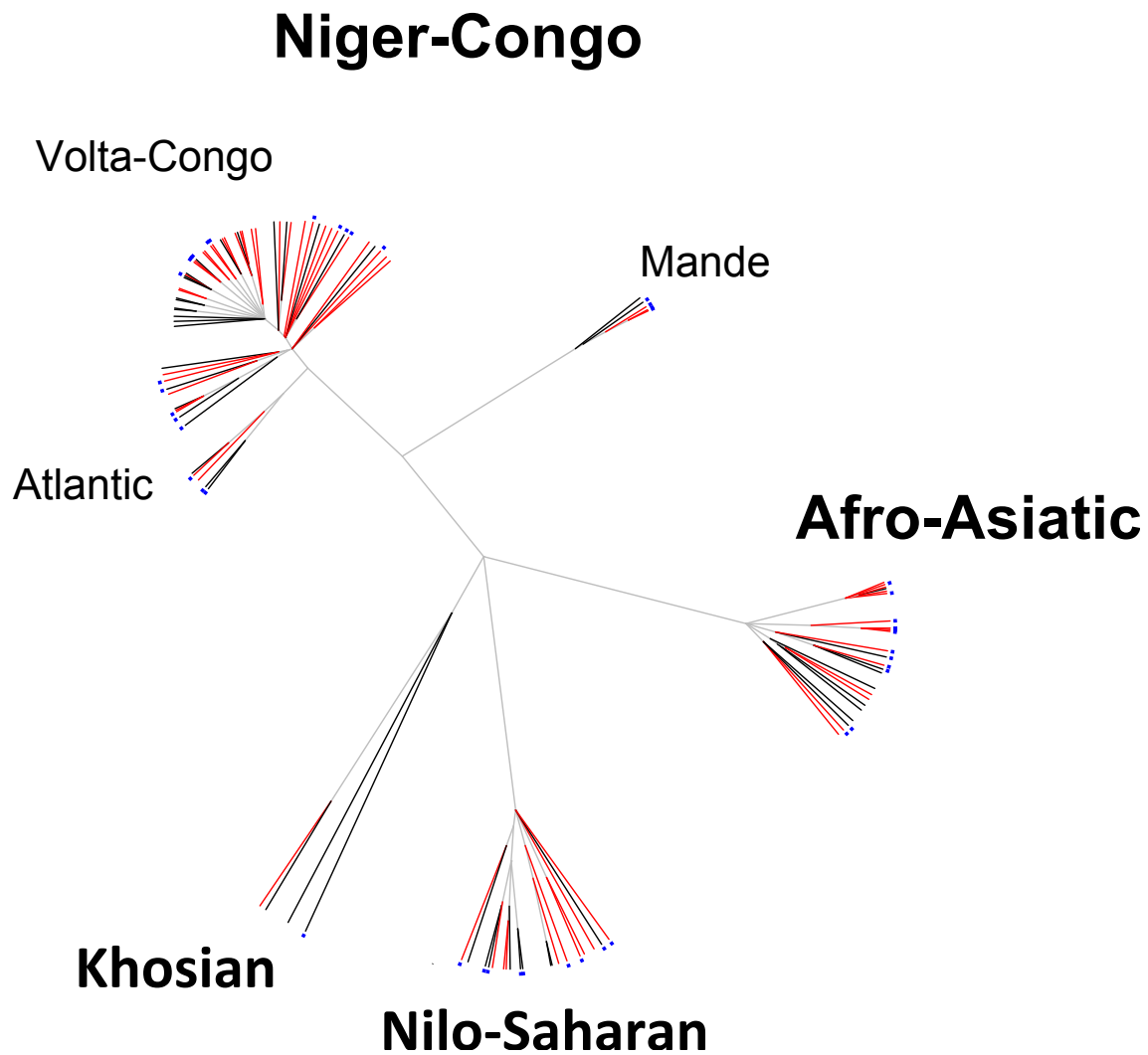
We begin our phylogenetic analysis with a hierarchical clustering of the selected cultural groups ( $N = 113$ ) in Africa [Murdock, 1969] according to the linguistic divergences postulated in Ethnologue [Lewis, 2009]. Following Walker et al [2012], we then utilize the estimated dates of lexical divergence produced by the ASJP (Automated Similarity Judgment Program) to define the prior expected proportionality of branch lengths [Holman et al, 2011]. Neither the clustering of languages in Ethnologue, nor the estimated times of language divergence produced by the ASJP are free of controversy [see peer commentary in Holman et al, 2011], with good reason — they are surely incorrect. There are two critical issues with the use of linguist data to infer the splitting of cultural groups: 1) linguistic evolution, like any form of cultural evolution, is not necessarily tree-like. There is, however, some tree-like structure to linguistic and cultural evolution, and we believe accounting for this structure is better than ignoring it completely [Gray et al, 2010]. And, 2) assumptions of constant rates of language evolution have been shown to be in conflict with empirical data [Gray et al, 2007]. ASJP estimates of divergence dates, however, do not assume constant evolutionary rates, and use empirical calibration on the basis of historical, archaeological, and other evidence, as suggested by Gray et al [2007].

Despite the fact that the phylogenetic tree used in our analysis is only of limited accuracy, in the Bayesian interpretation of the model fit in our analysis, phylogenetic information is utilized to construct prior beliefs concerning the expected covariance of the preferences for the FGMO trait across cultural groups due to shared ancestry. The fact that there is error in the ASJP estimates of divergence times is not necessarily a problem for our analysis. So long as the errors of estimated divergence times based on lexical similarity are roughly proportional across cultural groups, and do not vary as a function of stratification or FGMO prevalence, our inferences are unlikely to be biased (the actual calendar dates of linguistic divergences do not matter for our analysis).

Furthermore, we know of no other published phylogenies which estimate the divergence times (branch lengths) of all African languages in a unified framework; the ASJP phylogeny is also derived under more theoretical and empirical rigor than standard Bayesian approaches to phylogenetic reconstruction that utilize a simple binary coding procedure of cognate classes from Swadesh lists. Future studies will surely improve inference concerning the nature and dating of the somewhat tree-like, somewhat reticulated river-network-like, cultural evolutionary pasts of extant human groups [Towner et al, 2012]. When such data become available, our inferences herein should be re-investigated.

Lastly, it should be noted that our methods *do not* assume that the phylogenetic tree for FGMO is the same as the phylogenetic tree for language. This fact will be immediately apparent in the formulation of the mathematical details of our phylogenetic model. We go into even greater detail about the differences between our adaptive methods, and standard phylogenetic correlation methods in the Supplementary Materials, Appendix A Section 1.1.

Figure 2 displays our phylogeny with prior branch lengths scaled to the ASJP estimates. In the Supplementary Materials, Appendix A Section 1.3, we detail how exactly we bring the ASJP divergence times into our analysis, and how we integrate over uncertainty in these values.



**Fig. 2** Figure 2 displays the phylogeny of African languages used in this analysis. In this figure, branch lengths are presented as proportional to maximum likelihood AJSP divergence estimates; the actual phylogenetic tree implemented in the model, however, is a constrained random variable, allowing Bayesian integration over uncertainty in phylogenetic information. Black branches on the tree indicate non-stratified selection regimes. Red branches indicate stratified selection regimes. Gray branches indicate a mixture of stratified and non-stratified selection regimes. This mixture modeling allows for integration over uncertainty in deeper, non-terminal, branches. The blue points on the edge of the phylogenetic tree indicate the presence of practice of FGMo.

### 3.1.4 Modeling Adaptive Evolution

464

At each MCMC (Markov Chain Monte Carlo) iteration, a random painting of divergence times onto the branches of the phylogenetic tree is proposed. We then standardize the phylogeny to the unit interval such that  $t = 0$  is the time of the

465

466

467

deepest divergence and  $t = T = 1$  is the present. We then calculate two matrices,  $S$  and  $C$ , from the phylogeny. The  $S$  matrix is an  $N$  by  $N$  matrix ( $N$ =number of cultures included in this analysis=113) of the time points when cultural group  $n$  split from each and every other cultural group. The  $C$  matrix is an  $N$  by  $\Gamma$  matrix of the cut-points/transitions in epochs within a lineage, where  $\kappa(n)$  is the total number of epochs in lineage  $n$ , and  $\Gamma = \max(\kappa(n)) = 12$  is the maximum number of epochs in a cultural lineage observed in our data. A cultural lineage is defined as the path from tip of the phylogenetic tree to its most basal node. The history of the  $n^{th}$  lineage is then a series of  $\kappa(n)$  branch segments demarcated by epochs  $0 = t_n^0 < t_n^1 < t_n^2 < \dots < t_n^{\kappa(n)} = T$ , where each epoch constitutes a single kind of selective regime, to be defined more thoroughly later.

Equation 11 describes a stationary, Gaussian, and Markovian process with well defined moments; following, [Butler and King \[2004\]](#), we make a reducing assumption that in every lineage, evolution of a cultural trait,  $X$ , occurs along piecewise-constant selection regimes. Accordingly, the expected value of a trait evolving along a cultural lineage,  $n$ , can be defined as:

$$\mu_{[n]} = E[X_n(T)|X_n(0) = \beta_{[n,1]}] = e^{-\alpha T} \beta_{[n,1]} + \sum_{\gamma=2}^{\kappa(n)} e^{-\alpha T} (e^{-\alpha C_{[n,\gamma]}} - e^{-\alpha C_{[n,\gamma-1]}}) \beta_{[n,\gamma]} \quad (12)$$

where  $\beta_{[n,1:\kappa(n)]}$  is a parameter vector of length  $\kappa(n)$  painted with  $\theta$  parameters. The first cell of this vector,  $\beta_{[n,1]}$ , is defined to be equal to  $\theta_{anc}$ , the estimated trait value of the most basal node in the phylogeny, and all other cells are populated with differing  $\theta$  parameters that describe the hypothesized selective regime acting on the  $\gamma^{th}$  epoch in lineage  $n$ . The specific way that other  $\theta$  parameters are painted on to the phylogenetic tree will be made clearer when we describe the exact models being compared in this study.

Regarding covariance, we assume that when  $t < S_{[n,m]}$  lineages  $n$  and  $m$  evolved as a single group, and when  $t > S_{[n,m]}$  the two lineages evolve independently. Accordingly, the covariance matrix  $V_{[n,m]}$  can be defined as:

$$V_{[n,m]} = \text{Cov}[X_n(T), X_m(T)|X_n(0) = X_m(0) = \beta_{[n,1]}] = \frac{\sigma^2}{2\alpha} e^{-2\alpha(T-S_{[n,m]})} (1 - e^{-2\alpha S_{[n,m]}}) \quad (13)$$

To complete the basic model definition, we define regularizing priors on the  $\alpha$  and  $\sigma$  parameters, which concentrate prior probability density near zero. We considered both half-Gaussian and half-Cauchy [[Gelman, 2006](#)] priors (see Supplementary Materials, Appendix A Section 1.5 and 1.6 for more information on these priors). In our final analysis, we used the following half-Gaussian priors:

$$\alpha \sim \text{Normal}(0, 5)T[0, \infty] \quad (14)$$

$$\sigma \sim \text{Normal}(0, 5)T[0, \infty] \quad (15)$$

We use regularizing unit normal priors on each cell of the  $\theta$  parameter vector. These priors can be understood as imposing the Bayesian corollary of Tikhonov regularization [[Tikhonov and Arsenin, 1977](#)], or ridge regression [[Hoerl and Kennard, 1970](#)], and aid in the identification of  $\theta$  and  $\alpha$ , which are not necessary well identified otherwise:

$$\theta \sim \text{Normal}(0, 1) \quad (16)$$

We then model the data using a multivariate normal distribution parameterized to accept a Cholesky factor,  $L_V$ , from the decomposition of the variance-covariance matrix,  $V$ , in place of the variance-covariance matrix itself. This parameterization of the model improves the performance of the HMC estimation process for technical reasons that are outlined in the Stan manual [Stan Development Team, 2013b].

$$\Psi \sim \text{Multivariate Normal Cholesky}(\mu, L_V) \quad (17)$$

$\Psi$  is a parameter vector which represents the strength of evolving social preferences for FGMO, and:

$$FGMO_{[n]} \sim \text{Bernoulli}(\text{Logistic}(\Psi_{[n]})) \quad (18)$$

returns the log probability of the data,  $FGMO$ , conditioned on the estimated social preferences for FGMO;  $FGMO$  is a vector of binary indicators of the practice of FGMO.

### 3.1.5 Model Construction

In this paper, we fit two models to the data. In the first, we conceive of the evolutionary dynamics as an Ornstein-Uhlenbeck process with a single global optima for all cultural groups, an OU(1) model. This model serves as a null model, where stratification plays no role in the adaptive evolution of FGMO. In this case, we paint every branch of the phylogeny other than the most basal node,  $\theta_{anc}$ , with a single parameter that represents a single global selection regime,  $\theta_{gsr}$ .

In the second model, we conceive of the evolutionary dynamics as an Ornstein-Uhlenbeck process with separate optima for stratified and non-stratified cultural groups, an OU(2) model. In this case, we paint the phylogeny using two  $\theta$  parameters, with one corresponding to a stratified selection regime,  $\theta_{ssr}$ , and one corresponding to a non-stratified selection regime,  $\theta_{nssr}$ .

Deep ancestral branches are not easily classifiable as stratified or non-stratified, but the phylogenetic tree contains information on the probability of stratification in deeper branches conditional on the state of the branch tips and the strength of selection for stratification along the branches. Accordingly, at each MCMC iteration, we estimate the probability of stratification in deep branches, by using the Butler and King [2004] method to model the culture-group specific evolution of stratification using an OU(15) model (See Supplementary Materials, Appendix A Sections 1.2 and 1.6 for details and model diagnostics). Following this, we model the evolution of FGMO, conditioned on the estimated state of stratification at every node in the phylogeny.

More formally, in the OU(2) model, at each MCMC iteration, we run an adaptive phylogenetic model on stratification using the observed binary data variable *Strat*. We model:

$$\Psi_{strat} \sim \text{Multivariate Normal Cholesky}(\mu_{strat}, L_{V_{strat}}) \quad (19)$$

where  $\mu_{strat}$  and  $L_{V_{strat}}$  are corollaries to  $\mu$  and  $L_V$ , and are derived from equations 12 and 13 in the same way;  $\text{logistic}(\Psi_{strat})$  is a parameter vector which represents the log odds that a given cultural group is stratified, and:

$$Strat_{[n]} \sim \text{Bernoulli}(\text{Logistic}(\Psi_{strat[n]})) \quad (20)$$

provides the log probability of the observed stratification data conditional on the proposed parameter values.

The  $\beta_{strat}$  matrix for the stratification model is painted with  $\theta$  parameters such that the branches for each language family and subfamily have unique parameters (e.g. language family parameters are  $\theta_{Nilo-Saharan}$ ,  $\theta_{NigerCongo}$ , etc., and language subfamily parameters are  $\theta_{Omotoc}$ ,  $\theta_{Mande}$ , etc. ). The way in which these parameters are painted onto the branches is made clear in the Supplementary Stan code, using variables *GID2* and *GID3* from the Supplementary Data.

Under such a model it is straightforward to calculate the probability of stratification in any epoch in any lineage,  $A_{[n,\gamma]}$ , from Equation(12) as:

$$A_{[n,\gamma]} = \text{Logistic}(e^{-\alpha T} \beta_{strat[n,1]} + \sum_{\gamma=2}^{\kappa(n)} e^{-\alpha T} (e^{-\alpha C_{[n,\gamma]}} - e^{-\alpha C_{[n,\gamma-1]}}) \beta_{strat[n,\gamma]}) \quad (21)$$

To model the evolution of FGMO as a function of stratification, we then paint the tips of the phylogeny (in the FGMO model) with parameters using observed data:

$$\beta_{[n,\kappa(n)]} = \begin{cases} \theta_{ssr}, & \text{if } Strat_{[n]} = 1 \\ \theta_{nssr}, & \text{if } Strat_{[n]} = 0 \end{cases} \quad (22)$$

and paint all non-terminal epochs (e.g. where  $\gamma \neq \kappa(n)$ ) across all lineages (in the FGMO model) as a mixture of  $\theta_{ssr}$  and  $\theta_{nssr}$  with mixing proportions determined by  $A$ :

$$\beta_{[n,\gamma]} = \theta_{ssr} A_{[n,\gamma]} + \theta_{nssr} (1 - A_{[n,\gamma]}) \quad (23)$$

### 3.2 Software Environment

Each complete model was fit using the Stan 2.2.0 C++ library, using the No-U-Turn sampler, a variant of Hamiltonian Monte Carlo [Stan Development Team, 2013a]. We ran multiple small chains to ensure that the models were well defined and converged to similar posterior regions. Our final results are based on the results from a single long chain, run for 20,000 warm-up iterations and 400,000 sampling iterations, thinned at an interval of 20. We found that our model was well identified for all  $\theta$  parameters, for  $\alpha$ , and effectively identified for  $\sigma$ , although  $\sigma$ , itself is not numerically well identified. Convergence diagnostics and model identification are discussed in detail in the Supplementary Materials, Appendix A Sections 1.4 and 1.6.

All pre-processing of data and post-processing of MCMC samples was conducted using the R environment for statistical computing [R Core Team, 2013]. The package *RStan* was used to communicate between the C++ model and the R statistical environment.

#### 3.2.1 Model Comparison

We compare models using the Watanabe-Akaike information criterion (WAIC) [Gelman et al, 2013], which is a more fully-Bayesian generalization of the standard



Akaike information theoretic criteria, AIC. Computed WAIC is defined as:

$$\text{WAIC} = -2(lppd - p_E) \quad (24)$$

The computed log pointwise posterior predictive density,  $lppd$ , is defined as:

$$lppd = \sum_{n=1}^N \log\left(\frac{1}{Q} \sum_{q=1}^Q Pr(FGM_{o[n]}|\Psi_{[n,q]})\right) \quad (25)$$

where  $q = 1 \dots Q$  references the index of simulations from the posterior distribution. The effective number of parameters,  $p_E$  is computed as:

$$p_E = \sum_{n=1}^N \text{Var}_{q=1}^Q \log(Pr(FGM_{o[n]}|\Psi_{[n,q]})) \quad (26)$$

where the symbol  $\text{Var}_{q=1}^Q$  represents the function to calculate the sample variance over the posterior simulations.

## 4 Results

Analysis of the data with Ornstein-Uhlenbeck process models and information theoretic model comparison methods shows that while stratified cultural groups are on average slightly more likely to practice FGMO than non-stratified groups, there is only weak evidence that the stratification model provides a better fit for the data than the null model.

Table 1 presents the key parameters of our models, showing the posterior mean and medians estimates, as well as the central ninety-five percent posterior confidence intervals (95PCIs) from the fitted OU(1) and OU(2) models. Table 2 presents the results of model comparison using WAIC.

**Table 1** Key parameter estimates from our model. The mean and median are point estimates of the posterior distribution; the 2.5 and 97.5 percent equal tail posterior confidence intervals (PCI) present the dispersion of the posterior distribution. We note that  $\sigma$  dominates  $\alpha$  in both the OU(1) and OU(2) models, which is indicative that the distribution of FGMO in Africa is better explained by drift and/or selective forces operating orthogonally to stratification, than by stratification itself. In the the OU(2) model however, we find that stratification relative to non-stratification is weakly, but positively, associated with elevated social pressure for FGMO.

	Model	Mean	Median	2.5% PCI	97.5% PCI
$\theta_{anc}$	OU(2)	-0.075	-0.078	-1.974	1.853
$\theta_{ssr}$	OU(2)	-0.04	-0.042	-1.868	1.761
$\theta_{nssr}$	OU(2)	-0.278	-0.289	-2.106	1.56
$\alpha$	OU(2)	2.781	2.636	0.412	6.105
$\sigma$	OU(2)	6.891	6.712	2.234	12.502
$\theta_{anc}$	OU(1)	-0.047	-0.039	-1.994	1.889
$\theta_{gsr}$	OU(1)	-0.18	-0.212	-1.789	1.547
$\alpha$	OU(1)	2.652	2.496	0.337	5.88
$\sigma$	OU(1)	7.404	7.102	2.701	14.023

**Table 2** Results of formal model comparison using WAIC. WAIC,  $p_E$ , and lppd are defined in the text. The symbol dWAIC indicates the difference in WAIC between the best model and the second model, and the symbol wWAIC indicates the weight in probability that the specified model will make the best predictions on new data, conditional on the set of models being considered. We note that the OU(1) model outperforms the OU(2) model. This result indicates that the distribution of FGMo across our sample of African cultural groups can be most parsimoniously explained by a drift model with a single global optima, as opposed to a model with separate optima for stratified and non-stratified societies. However, these wWAIC values are very close, and the application of WAIC to these model is not completely justified given the relative strength of drift; as such, we argue that both models are important to consider. See Supplementary Materials, Appendix A Section 1.5 for a discussion of methodological issues that arise with WAIC, or any information criterion, as  $\sigma$  increases relative to  $\alpha$  in adaptive phylogenetic models.

Model	$p_E$	lppd	WAIC	dWAIC	wWAIC
OU(1)-Null	28.099	-28.409	113.017	0	0.848
OU(2)-FixedBranchTips	27.407	-30.821	116.458	3.441	0.152

Inspection of the  $\theta_{ssr}$  and  $\theta_{nssr}$  parameter estimates from the OU(2) model in Table 1, shows that there is a moderate difference in the mean estimated strength of social preferences for FGMo as a function of stratification. Converting these parameter estimates to the probability scale, we find that stratified cultural groups are about 6 percent more likely, on average, to practice FGMo than non-stratified cultural groups. The confidence intervals on these estimates are wide and largely overlap one another, however. This being said, estimates of  $\theta_{ssr}$  and  $\theta_{nssr}$  are in the direction predicted by our evolutionary models.

Model comparison with WAIC, however, shows that the OU(1) model is preferred over the OU(2) model. So while increased stratification appears to be weakly associated with increased odds of FGMo, accounting for stratification does not improve predictions sufficiently to compensate for the increased model complexity in an information theoretic framework.

In conclusion, although we present evidence that social stratification places positive selection pressure on social preferences for FGMo, it is evident that  $\sigma$  (which accounts for drift and/or selection dynamics operating orthogonally to stratification) plays a very important role in explaining the cross cultural distribution of FGMo. Future research is needed to disentangle the effects of drift from other possible selective drivers of FGMo.

## 5 Discussion

Historical records, albeit largely speculative, that place the origins of FGMo in ancient Egypt, or the strongly Egyptian-influenced and fabulously wealthy trading city of Meroe, are consistent with our model-based findings that variance in male mating value and social stratification should be causally linked to the *de novo* origins of FGMo, since ancient Egyptian society was marked by complex stratification [Murdock, 1957]. Further, the results of our adaptive phylogenetic analysis suggest that stratification is associated, though only weakly, with selection pressure on the uptake of FGMo. However, the fact that  $\sigma$  dominates  $\alpha$  in explaining the evolutionary dynamics of the FGMo trait in Africa indicates that drift and/or selective forces operating orthogonally to social stratification play a

very significant role in explaining the current distribution of the FGMo trait. This is entirely consistent with the view of earlier scholars who suggest that FGMo became decoupled from the signaling of marriageability and chastity as it spread into less stratified populations [Mackie, 2000].

While our model-based results indicate that stratification is needed for the origins of FGMo in a single culture, once the trait has arisen in a single cultural group, its adoption in any subsequent cultural groups does not require stratification. Once FGMo has arisen in a cultural group, and the system has equilibrated such that  $s$ —a measure of frequency-dependent, conformist, and prestige biases [Boyd and Richerson, 1985; Henrich and Gil-White, 2001; Henrich and Boyd, 1998]—functions to increase pressure on individuals to undergo FGMo, the genesis of FGMo in subsequent groups can occur strictly as a function of  $s$ , irrespective of stratification. Thus, the dynamics of the *de novo* origins of FGMo differ from the subsequent dynamics of inter-cultural transmission. Though our adaptive phylogenetic model illustrates that the evolution of FGMo is not strictly dependent on stratification, the model still suggests that stratification increases the likelihood of inter-group transmission or adoption of FGMo.

In our theoretical models, we focused mainly on differentiating the evolutionary forces responsible for the *de novo* origins of FGMo and its subsequent maintenance in a population. The results of our phylogenetic analysis, however, show that a holistic characterization of the cross-cultural origins of FGMo requires more explicit treatment of the dynamics underlying inter-group transmission as well.

The nature of the relationship between  $s$  and the inter-cultural transmission of FGMo is complex and will be structured by culturally and historically particular processes. For instance, imagine that migrants from a non-FGMo culture enter an FGMo cultural area. In such a case,  $s$  and conformist biases now act to place selective pressure on the uptake of FGMo by these inter-cultural migrants, even if stratification was not present in the migrants' cultural group. Such an effect would be exacerbated if there were perceived prestige differences between the FGMo practicing cultural group and the non-FGMo practicing migrant group that cause migrants to copy the behaviors of the prestigious group. This pattern would be likely if FGMo is associated with the prestigious class in stratified areas.

A similar pattern would be observed if adoption of the cultural traits of a dominant or prestigious cultural group is an essential step in successfully integrating into that culture; such dynamics might partially underlie the observation by Cronk [2004] that Kenyan Mukogodo appear to have adopted FGMo to hasten their transition to becoming Maasai, or the observation by Ericksen [1989] that the Fur appear to have adopted the trait to marry with the nomadic stock-owning and more wealthy Zaghawa neighbours. Similarly non-practicing displaced ethnic communities in Darfur (Sudan) have adopted FGMo when moving to cities, just as in the state of Khartoum where migrants from West Africa now cut their daughters to gain acceptance in their new host community [DFID, 2013].

If successful inter-cultural migrants are emulated or held in high esteem, or if migration events involve a large portion of a cultural group, then frequency-dependent, conformist, and prestige biases can result in the propagation of FGMo back into the migrants' natal cultural group. Since our phylogenetic analysis considered FGMo as present in any cultural group in which a single ethnographer reported any normative occurrence of FGMo, our methods are sensitive to detecting such transmission events, even if the FGMo rate in the migrants' cultural

group is actually quite low. A different classification scheme (for example, only classifying cultures with high, >90% prevalence rates of FGMO) might lead to a stronger correlation of FGMO and social stratification, but might miss detecting the nuanced processes which drive the social transmission of costly traits like FGMO.

Different patterns of transmission are also possible. We know, for example, that during the intense military turmoil of the mid 18th century in the Great Lakes region of central Africa that refugee segments of Bantu speaking Abaluhya lineages, subsequently known as the Tiriki, adopted the age set organization of their Nilotic Terik neighbors. In return for asylum with the more military-prepared Terik, Tiriki elders accepted the full set of initiation rituals for their sons, including circumcision and seclusion, in a bid to obtain protection [Levine and Sangree, 1962; Boyd et al, 1997].

Contrarily, it is conceivable that in many contexts the practice of FGMO marks an ethnic boundary [see McElreath et al, 2003] between groups that practice FGMO and those that do not. In such contexts,  $s$  could militate against adoption of the cultural traits of out-groups, and work against the inter-cultural transmission of FGMO to neighboring population. Case studies in which such socially or behaviorally marked ethnic boundaries function to prevent the transmission of information and behavior across groups are well described in the literature [Barth, 1998; Van den Berghe, 1987], although their pertinence to FGMO has been questioned [Mackie, 2000].

As such, we expect that the dynamics governing the transmission of FGMO and other costly traits will be historically and locally contingent based on the nature of inter-cultural contact. We suspect that the mechanisms underlying these inter-cultural dynamics are likely to include frequency-dependence or conformist learning biases, as outlined in our introductory review of how FGO is embedded in deeply-held societal conventions, as well as prestige-biased learning, whereby adopting FGMO may facilitate inclusion in a more powerful or prestigious population. We also expect these dynamics to be differentially visible to different methodological approaches concerning classification and prevalence of the FGMO trait. Similarly we note that the decline in FGMO is particularly striking in countries such as Kenya and Tanzania, where FGMO is less prevalent across constituent ethnic groups [UNFPA-UNICEF, 2013], suggesting that abandonment is more tolerable, even practicable, with non-practitioners as neighbors. Future research should investigate in more detail the socio-ecological circumstances that aid in the inter-cultural transmission of FGMO and those that hinder transmission and hasten abandonment.

Our conclusions are generally consistent with those of Murdock [1959], who argues that the trait of FGMO—in his terminology, excision—largely spread across Africa through cultural diffusion. For example, in East Africa, southern Nilotes and Bantu are thought to have adopted the custom from Cushitic neighbors, given that there is little evidence of FGMO among other Nilotes and the apparent Cushitic roots to the linguistic term used for the operation [Murray, 1979]. In West Africa too there is strong inferential evidence that FGMO was adopted by some of the neighbors of the FGMO practicing Mande, for example the Kissi and the Kran, but not the Kpelle, Guro and Gbande, but the details of the adoption during this historic period of diffusion are unclear [Erickson, 1989]. Furthermore ethnographic reports [Stannus, 1919] indicate that the Yao of northwestern Mozambique and

southern Tanzania appear to have adopted female ‘circumcision’ during their close collaboration with Swahili and Arab slave traders in the 19th century, although now it is largely dropped. We note, however, that these conclusions regarding the prominence of historic diffusion in the distribution of the trait are based largely on the observation that FGMO appears in populations that do not necessarily share a common language, geography, and cultural history, and that our methods [like those of [Towner et al, 2012](#)] provide a significant improvement in inference.

## 5.1 Implications for Policy

Here we sidestep ethical arguments regarding whether or not FGMO should be abandoned, as well as the debates over who has moral authority to take the lead in such initiatives, recognizing that this can only be locally adjudicated. As [Shell-Duncan and Hernlund \[2000, p126\]](#) observe for the Rendille of Kenya, “awareness of the fact that female ‘circumcision’ is associated with adverse health consequences is widespread, yet the Rendille view the risks as worth taking in light of the implications for marriageability.” It is quite clear that there are social, psychological and physical consequences to the practice of genital cutting, and that these need to be weighed very carefully in each case. But, if abandonment is viewed as appropriate, there remains the question of how to do achieve this. Many strategies are currently discussed, including medicalization of the practice (for example, the use of less extreme procedures in hospitals by specially trained practitioners), mass education campaigns, formal legislation with criminalization of operators and their clients, withdrawing of foreign aid programs, developing ritual alternatives, or simply relying on the processes of development and the erosion of traditional culture [[Shell-Duncan and Hernlund, 2000](#)].

Many scholars argue that knowledge of the origins of FGMO will not contribute to a determination of its current significance, others maintain that understanding the historic roots of the tradition helps justify the persistence of the custom to disapproving outsiders, while yet others insist that the origins and maintenance of the practice cannot be conceptually separated from development of strategies for its elimination [[Mackie, 2000](#)]. The present analysis, with its implications for the importance of frequency dependent biases in the maintenance of FGMO, support this third viewpoint. Specifically they point to the potential value of programs that foster contracts within small communities whereby all parties make a pact not to send their daughters for the operation. [Mackie \[1996\]](#) has made precisely this point, drawing an analogy between successful campaigns to eradicate foot-binding in China through fostering of such local contracts. Furthermore, the most recent evaluation of the status of FGMO globally suggests that public declarations, or collective announcements of abandonment by village delegates, have for the most part been very successful in supporting change [[UNFPA-UNICEF, 2013](#)]. There is nevertheless huge variation within and between communities in the rate of abandoning [[Shell-Duncan and Hernlund, 2007](#)], which begs for further study.

**Acknowledgements** Author Contributions:CTR designed and tested the Bayesian models, conducted analysis, and wrote/edited the paper. PS designed the mathematical models and wrote/edited the paper. KE collected and compiled the ethnographic information and edited the paper. PL conducted the alternative phylogenetic analysis and wrote/edited the paper. MBM conceived the study, reviewed the literature, and wrote/edited the paper.

We thank the UC Davis Behavioral Ecology and Cultural Evolution lab groups for helpful comments and critiques, Richard McElreath and Andrew Gelman for providing code and advice on model comparison using WAIC, and the Stan development team for making Stan freely available and open source, and for providing impressive levels of software support and consulting. Dr. Mark Grote gave significant advice concerning model notation and provided statistical consulting that much improved our methodological framework.

## References

- Anderson S (2003) Why dowry payments declined with modernization in europe but are rising in india. *Journal of Political Economy* 111(2):269–310
- Banks E, Meirik O, Farley T, Akande O, Bathija H, Ali M (2006) Female genital mutilation and obstetric outcome: Who collaborative prospective study in six african countries. *Lancet* 367(9525):1835–1841
- Barth F (1998) *Ethnic groups and boundaries: The social organization of culture difference*. Waveland Press
- Becker GS (1981) *A Treatise on the Family*. Harvard University Press
- Bell D, Song S (1994) Explaining the level of bridewealth. *Current Anthropology* 35(3):311–316
- Van den Berghe PL (1987) *The ethnic phenomenon*. ABC-CLIO
- Borgerhoff Mulder M (1995) Bridewealth and its correlates: quantifying changes over time. *Current Anthropology* 36(4):573–603
- Borgerhoff Mulder M, Nunn CL, Towner MC (2006) Cultural macroevolution and the transmission of traits. *Evolutionary Anthropology: Issues, News, and Reviews* 15(2):52–64
- Boyd R, Richerson PJ (1985) *Culture and the Evolutionary Process*. University of Chicago Press
- Boyd R, Borgerhoff Mulder M, Durham WH, Richerson PJ (1997) Are cultural phylogenies possible. *Human by Nature: Between biology and the social sciences* pp 355–384
- Butler MA, King AA (2004) Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *The American Naturalist* 164(6):683–695
- Chiappori PA, Fortin B, Lacroix G (2002) Marriage market, divorce legislation, and household labor supply. *Journal of Political Economy* 110(1):37–72
- Cronk L (2004) *From Mukogodo to Maasai: Ethnicity and cultural change in Kenya*. Westview Pr
- Daniels RE (1970) *By rites a man: A study of the societal and individual foundations of tribal identity among the kipsigis of kenya*. PhD thesis, University of Chicago
- Davison J (1996) *Voices from Mutira: Changes in the Lives of Rural Gikuyo Women, 1910-1995*. Lynne Rienner Publishers
- DFID (2013) Department for international development: Sudan free of female genital cutting, gb-1-203407. URL <http://devtracker.dfid.gov.uk/projects/GB-1-203407/>
- Dickemann M (1979) The ecology of mating systems in hypergynous dowry societies. *Social Science Information* 18(2):163–195
- Dorkenoo E (1994) *Cutting the rose: female genital mutilation: the practice and its prevention*. Minority Rights Publications London

- Ericksen KP (1989) Female genital mutilations in africa. *Cross-Cultural Research* 23(1-4):182–204
- Gaulin SJ, Boster JS (1990) Dowry as female competition. *American Anthropologist* 92(4):994–1005
- Gelman A (2006) Prior distributions for variance parameters in hierarchical models (comment on article by browne and draper). *Bayesian analysis* 1(3):515–534
- Gelman A, Hwang J, Vehtari A (2013) Understanding predictive information criteria for bayesian models. *Statistics and Computing* pp 1–20, DOI 10.1007/s11222-013-9416-2
- Grafen A (1990) Biological signals as handicaps. *Journal of Theoretical Biology* 144(4):517–546
- Gray RD, Greenhill SJ, Ross RM (2007) The pleasures and perils of darwinizing culture (with phylogenies). *Biological Theory* 2:4
- Gray RD, Bryant D, Greenhill SJ (2010) On the shape and fabric of human history. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365(1559):3923–3933
- Hayden B, et al (1997) Prehistoric rites of passage: A comparative study of transegalitarian hunter-gatherers. *Journal of Anthropological Archaeology* 16(2):121–161
- Hayes RO (1975) female genital mutilation, fertility control, women’s roles, and the patrilineage in modern sudan: a functional analysis1. *American Ethnologist* 2(4):617–633
- Henrich J, Boyd R (1998) The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior* 19(4):215–241
- Henrich J, Gil-White FJ (2001) The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior* 22(3):165–196
- Hoerl AE, Kennard RW (1970) Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* 12(1):55–67
- Hoffman MD, Gelman A (2013) The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research in press*
- Holman EW, Brown CH, Wichmann S, Müller A, Velupillai V, Hammarström H, Sauppe S, Jung H, Bakker D, Brown P, et al (2011) Automated dating of the worlds language families based on lexical similarity. *Current Anthropology* 52(6):841–875
- Leonard L (2000) Adopting female “circumcision” in southern chad: the experience of myabe. *Female “circumcision” in Africa: culture, controversy, and change* Boulder, Colorado, Lynne Rienner pp 167–192
- Levine RA, Sangree WH (1962) The diffusion of age-group organization in east africa: A controlled comparison. *Africa* 32(02):97–110
- Lewis MP (2009) *Ethnologue: Languages of the world* sixteenth edition. Dallas, Tex: SIL International Online version: <http://www.ethnologue.com>
- Little CM (2003) Female genital circumcision: medical and cultural considerations. *Journal of Cultural Diversity* 10(1):30
- Mackie G (1996) Ending footbinding and infibulation: A convention account. *American Sociological Review* pp 999–1017



- Mackie G (2000) Female genital cutting: the beginning of the end. Female “circumcision” in Africa: culture, controversy, and change Boulder, Colorado, Lynne Rienner pp 253–282
- Mackie G (2003) Female genital cutting: a harmless practice? *Medical Anthropology Quarterly* 17(2):135–158
- Mackie G, LeJeune J (2009) Social dynamics of abandonment of harmful practices: A new look at the theory. Tech. rep., UNICEF Innocenti Research Centre
- McElreath R, Boyd R, Richerson PJ (2003) Shared norms and the evolution of ethnic markers. *Current Anthropology* 44(1):pp. 122–130
- Murdock GP (1957) World ethnographic sample. *American Anthropologist* 59(4):664–687
- Murdock GP (1959) Africa: Its peoples and their culture history
- Murdock GP (1968) World sampling provinces. *Ethnology* 7(3):305–326
- Murdock GP (1969) *Ethnographic atlas*. University of Pittsburgh Press
- Murray JM (1979) *The Kikuyu Female Circumcision Controversy, with Special Reference to the Church Missionary Society’s “sphere of Influence”*. University Microfilms
- Myers RA, Omorodion FI, Isenalumhe AE, Akenzua GI (1985) Circumcision: its nature and practice among some ethnic groups in southern nigeria. *Social Science & Medicine* 21(5):581–588
- Pagel M (1994) Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society of London Series B: Biological Sciences* 255(1342):37–45
- Pagel M, Meade A (2006) Bayesian analysis of correlated evolution of discrete characters by reversible-jump markov chain monte carlo. *The American Naturalist* 167(6):808–825
- Paige KE, Paige JM (1981) *The politics of reproductive ritual*. University of California Press
- Quisumbing AR, Maluccio JA, et al (2000) *Intrahousehold allocation and gender relations: New empirical evidence from four developing countries*. International Food Policy Research Institute Washington, DC
- R Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, URL <http://www.R-project.org/>, ISBN 3-900051-07-0
- Shell-Duncan B, Hernlund Y (2000) *Female circumcision in Africa: Culture, Controversy, and Change*. Lynne Rienner Publishers
- Shell-Duncan B, Hernlund Y (2007) Are there “stages of change” in the practice of female genital cutting? qualitative research finding from senegal and the gambia. *African Journal of Reproductive Health* 10(2):57–71
- Silverman EK (2004) Anthropology and circumcision. *Annual Review of Anthropology* pp 419–445
- Smith EA, Bird RB, Bird DW (2003) The benefits of costly signaling: Meriam turtle hunters. *Behavioral Ecology* 14(1):116–126
- Stan Development Team (2013a) Stan: A c++ library for probability and sampling, version 2.0. URL <http://mc-stan.org/>
- Stan Development Team (2013b) Stan Modeling Language Users Guide and Reference Manual, Version 2.0. URL <http://mc-stan.org/>
- Stannus H (1919) *The Wayao of Nyasaland*. Mass



- Tikhonov A, Arsenin VY (1977) Methods for solving ill-posed problems. John Wiley and Sons, Inc 909
- Toubia N, Izett S (1998) Female genital mutilation: An overview. World Health Organization Geneva, Switzerland 910
- Towner MC, Grote MN, Venti J, Borgerhoff Mulder M (2012) Cultural macroevolution on neighbor graphs. *Human Nature* 23(3):283–305 911
- UNFPA-UNICEF (2013) Unfpa-unicef joint evaluation of the unfpa-unicef joint programme on female genital mutilation/cutting (fgm/c): Accelerating change. URL <http://www.unfpa.org/public/home/about/Evaluation/EBIER/TE/pid/10103> 912
- Van Gennep A (1960) The rites of passage, vol 3. Routledge 913
- Walker RS, Wichmann S, Mailund T, Atkisson CJ (2012) Cultural phylogenetics of the tupi language family in lowland south america. *PloS one* 7(4):e35,025 914
- Warzazi HE (1986) Report of the working group on traditional practices affecting the health of women and children. 915
- Whitehorn J, Ayonrinde O, Maingay S (2002) Female genital mutilation: Cultural and psychological implications. *Sexual and Relationship Therapy* 17(2):161–170 916