

## Supplementary Materials for

### High rate of extrapair paternity in a human population demonstrates diversity in human reproductive strategies

B. A. Scelza\*, S. P. Prall, N. Swinford, S. Gopalan, E. G. Atkinson, R. McElreath, J. Sheehama, B. M. Henn

\*Corresponding author. Email: bscelza@gmail.com

Published 19 February 2020, *Sci. Adv.* **6**, eaay6195 (2020)  
DOI: 10.1126/sciadv.aay6195

#### This PDF file includes:

Supplementary Materials and Methods

Fig. S1. Density plot of EPP Perception Task.

Fig. S2. Comparison of age for men in our paternity sample ( $n = 47$ ) with those not living in Omuhonga but not included in our sample ( $n = 68$ ).

Fig. S3. Comparison of age for women in our paternity sample ( $n = 59$ ) with those not living in Omuhonga but not included in our sample ( $n = 129$ ).

Fig. S4. Comparison of parity for women in our paternity sample ( $n = 59$ ) with those not living in Omuhonga but not included in our sample ( $n = 129$ ).

Fig. S5. Frequency distribution showing number of children in the sample, by mother.

Fig. S6. Frequency distribution showing number of children in the sample, by social father.

Fig. S7. Number of sampled children, by mother's age.

Fig. S8. Trace plots (top) and rank histograms (bottom) for model clustering mother and mother-husband dyad.

Fig. S9. Flowchart of DNA sample quality control, beginning with total number of population samples and the final number of samples available for paternity analysis.

# Supplementary Materials and Methods

## A. Marital and Reproductive History Interviews

For each individual, marital and reproductive histories were recorded. For the marital histories, prompts included questions like: At what age were you first married? Are you still married to that person? How many times have you been married? What is the current status of each marriage? Serial marriages are common, as is polygyny, and concurrent partnerships. While concurrency with informal partners is common for both men and women, there are clear distinctions between marital and non-marital partners, with marital partners being primary. Himba have formal marriage ceremonies, brideprice is paid, patrilian membership switches upon marriage, etc. Therefore, there is no ambiguity about who is a spouse, and the social father of children born into that marriage.

Reproductive histories were conducted slightly differently for men and women. Women were asked to list each birth and to denote whether that birth was pre-marital, marital or inter-marital (between marriages). Next women were asked about each marital birth, and within which marriage it occurred (if the woman had been married more than once). Men were asked to list the children of each of their wives separately. They were also asked to list the number of children they believed they had fathered outside of marriage, both before their first marriage, and subsequently. In more recent interviews, men grouped their non-marital children by their mother (e.g. two from girlfriend 1 and 1 from girlfriend 2).

Many of the informants for this study were interviewed multiple times between 2010 and 2017. In those cases, reproductive and marital histories were re-checked and updated when necessary. This included cases where a child may have previously been recorded as omoka and then switched, but much more commonly, where a person previously listed a child as being from the husband, and later asserted that the child was actually omoka. We chose to use the more recent account, as previous demographic work from long-term studies has shown that people are increasingly accurate the better they know a fieldworker. This accords with the responses we got when reporting back results to the community (that people were more likely to misreport in this direction) and with the paternity confidence data, which showed almost all cases of error being in this direction. This also explains the difference between the asserted rate of EPP in Scelza (2011) and the current results, as the former were from only the initial round of interviews conducted in 2010.

## B. Study Population Demographics

Men in the sample were older than women and their parity was higher (Table 1, main document). This can be attributed to several factors. First, men get married considerably later than women, and often their first wife is pre-menarchal, and so several years might pass until either their wife becomes fecund, or they take a second wife, who is typically closer to them in age. Parity is likely higher due to some skew in who was sampled (average parity should be the same for men and women due to the Fisher Condition). Men who had low parity (either overall or because they had yet to father any children within marriage) would not be included in our paternity analysis because only those having at least one child born to a wife were included in the paternity analysis. Men in their twenties and early thirties are more likely to fall into this category. Women on the other hand, due to earlier age at first birth and earlier age at marriage were more likely to be included across age categories. Another major factor is that parity for men was calculated as the total number of children born into their marriages. EPP cases were not subtracted out. This also artificially

inflates men’s number of children, from a biological perspective. Mean number of marriages was higher for men due to polygyny, and polygyny also leads to a higher percent currently married.

### C. Comparison of Sampled and Unsampled Populations

In order to assess how representative the mother-father-child triads included in our paternity study were, we have compared men and women who were included in this study with men and women with matching inclusion criteria in our larger demographic database, which includes almost the entire community of Omuhonga. For men, the inclusion criteria were that they must have been married at least once and had at least one child born within marriage (e.g. they are the social father of at least one child). We did not include men whose only record of social fatherhood was through unconsummated child marriage (see Methods). For women, they must have been in at least one consummated marriage and have born at least one child within marriage.

The men in our sample were representative of the greater Omuhonga community. A whole-network study of Omuhonga was conducted in 2016 and included 40 households, of which 34 had male heads of household. Of those 25 (74%) were included in our paternity study. Of the remaining nine, four provided DNA samples but they either did not have a child sampled, or their sample did not produce enough DNA to be included in the analysis. Married men who have children born to their wives tend to be older than the general male population. Modal age at marriage is 27, and with the high frequency of child marriage, men tend to be in their thirties before they begin having children with their wives. The age distribution of men in our paternity sample was very similar to the distribution of married men with children not in our sample, further indicating our data are representative of the community at large (Figure S2).

Women in the sample were also largely representative of the Omuhonga population at large. The women in the sample represented 29 of the 40 compounds in Omuhonga. Mean age of women in the sample was 44.8 years, compared to 43.6 years among unsampled women meeting the inclusion criteria (Figure S3). Parity in the sampled population was slightly higher for sampled women (6.6) compared to those unsampled (5.0), likely due to the increased chance of capturing a child in the sample the more children a woman has (Figure S4).

### D. Within Sample Variability in the Rate of EPP

To determine whether the EPP rate we found was the result of skew, where a few individuals with many children were driving the result, we looked at the distribution of children across mothers and social fathers in the sample. Figures S5 and S6 show that most of the parents in the sample have only one child represented. However, the older a woman is, the more children she is likely to have represented in the sample, due to the natural correlation between age and parity (Figure S7).

### E. Multilevel Models of EPP

The sample comprises mothers with varying numbers of both children and husbands. In principle, each mother and mother-husband pair could have a unique extra-pair paternity (EPP) rate. Estimates of both the average rate and the variation among mothers and mother-husband pairs can be constructed with a multilevel analysis. The following explains how the models represented in Figure 2 of the main manuscript were constructed.

#### i. Statistical model

Let  $y_i \in \{0, 1\}$  be the EPP status of child  $i$ . Each child has a corresponding mother with index  $mom[i]$  and a dyad comprising the mother and the child’s social father with index  $dyad[i]$ . Then the probability of  $y_i$  is given by:

$$y_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \alpha + z_{\text{mom}[i]}\sigma + x_{\text{dyad}[i]}\tau$$

where  $\alpha$  is the population average, the vector  $z$  contains standardized varying effects for each mother,  $\sigma$  is the standard deviation among mothers,  $x$  contains standardized varying effects for each dyad, and  $\tau$  is the standard deviation among dyads.

The model uses a non-centered parameterization, but is in fact an ordinary multi-level model. The priors are:

$$\alpha \sim \text{Normal}(0, 1.5)$$

$$z_j \sim \text{Normal}(0, 1)$$

$$x_k \sim \text{Normal}(0, 1)$$

$$\sigma \sim \text{HalfNormal}(0, 1)$$

$$\tau \sim \text{HalfNormal}(0, 1)$$

The  $\text{Normal}(0, 1.5)$  prior for  $\alpha$  establishes a uniform prior on the probability scale and  $\text{HalfNormal}(0, 1)$  indicates a normal distribution defined only on positive reals.

## ii. Measurement error

A complication in this analysis is that there is as much as 5% false assignment of paternity. Failing to account for this would artificially inflate our estimate of EPP. We incorporate false-positives into our model by using a mixture likelihood. Given false-positive probability  $f$ , the probability of observing a paternity assignment  $y_i = 1$  is:

$$\Pr(y_i = 1|p_i) = (1 - f)p_i + f(1 - p_i)$$

Since false negatives are extremely unlikely in our analysis, the probability of observing  $y_i = 0$  remains unchanged:

$$\Pr(y_i = 0|p_i) = 1 - p_i$$

## iii. Implementation

We implemented the model in Stan 2.19, a library for Bayesian model construction and fitting. We drew 8000 samples from 4 chains. All R-hat values were less than 1.01, and there were no divergent transitions or other signs of biased posterior exploration. We show the trace plots for the hyper-parameters and a selection of the varying effects in Figure S8. We also show the same traces as rank histograms (<https://arxiv.org/abs/1903.08008>), which make it much easier to spot problems with chains. Healthy chains, like these, should be relatively uniform.

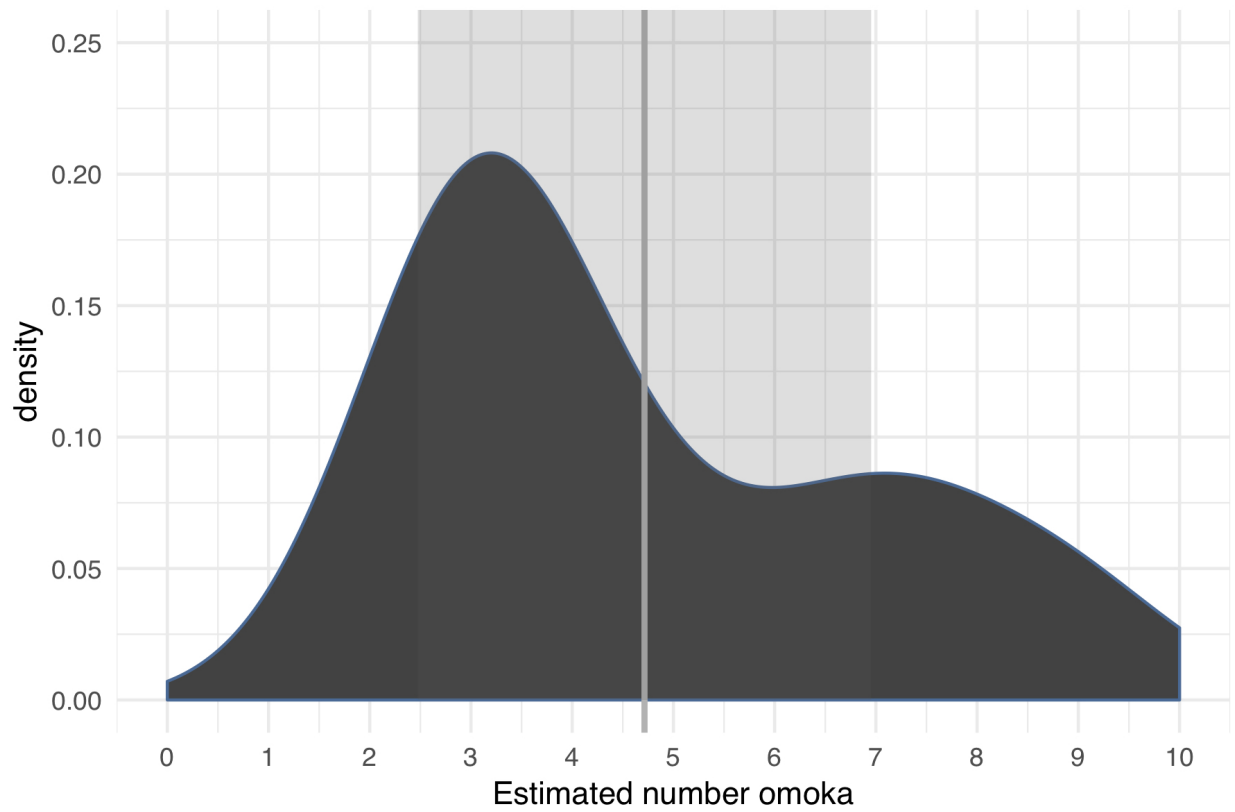
The marginal posterior distributions for  $\alpha$ ,  $\sigma$ , and  $\tau$  are summarized below.

	mean	sd	5.5%	94.5%	n_eff	Rhat
alpha	-0.07	0.36	-0.63	0.51	6422	1
sigma	0.75	0.55	0.07	1.75	1216	1
tau	2.01	0.54	1.18	2.89	2325	1

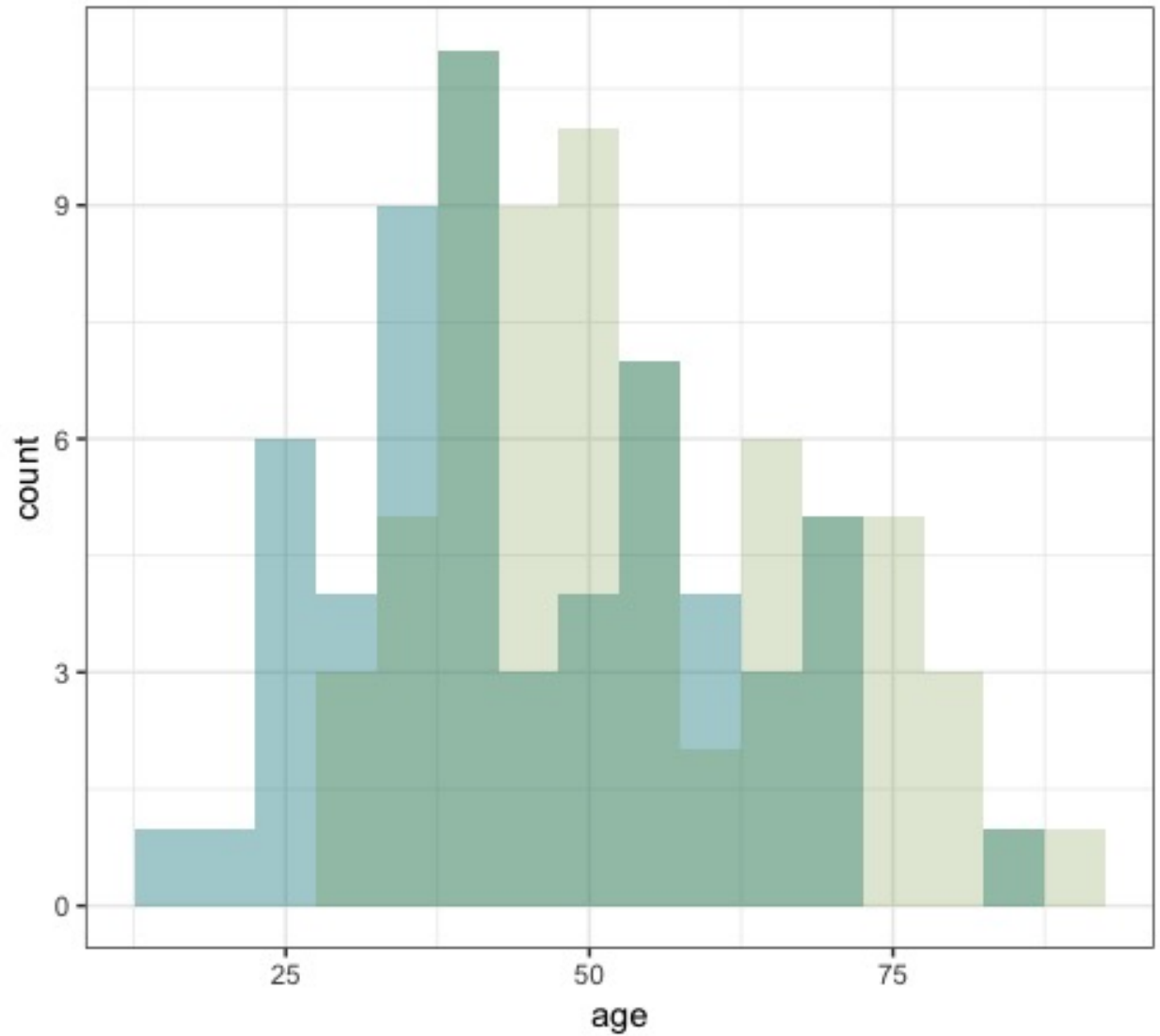
On the probability scale,  $\alpha$  implies an average posterior EPP rate with mean 0.48 and 95% credible interval from 0.32 to 0.66. Note that there is more variation among dyads ( $\tau$ ) than among mothers ( $\sigma$ ). This implies that features of specific pairings explain EPP variation more than features of specific women.

The posterior distribution EPP rate is plotted in Figure 2 in red, against both the prior (dashed) and the naive posterior rate from ignoring clustering by mothers and mother-father dyads (black). We also considered models with no clustering (the black density in figure 2), with clustering only by mother, and clustering only by dyad. All models with clustering produce similar inferences. Code for each, together with WAIC model comparison, can be found in the script at [https://github.com/rmcelreath/Himba\\_EPP](https://github.com/rmcelreath/Himba_EPP). For comparison, models which do not include potential false-positives are shown in lighter colors.

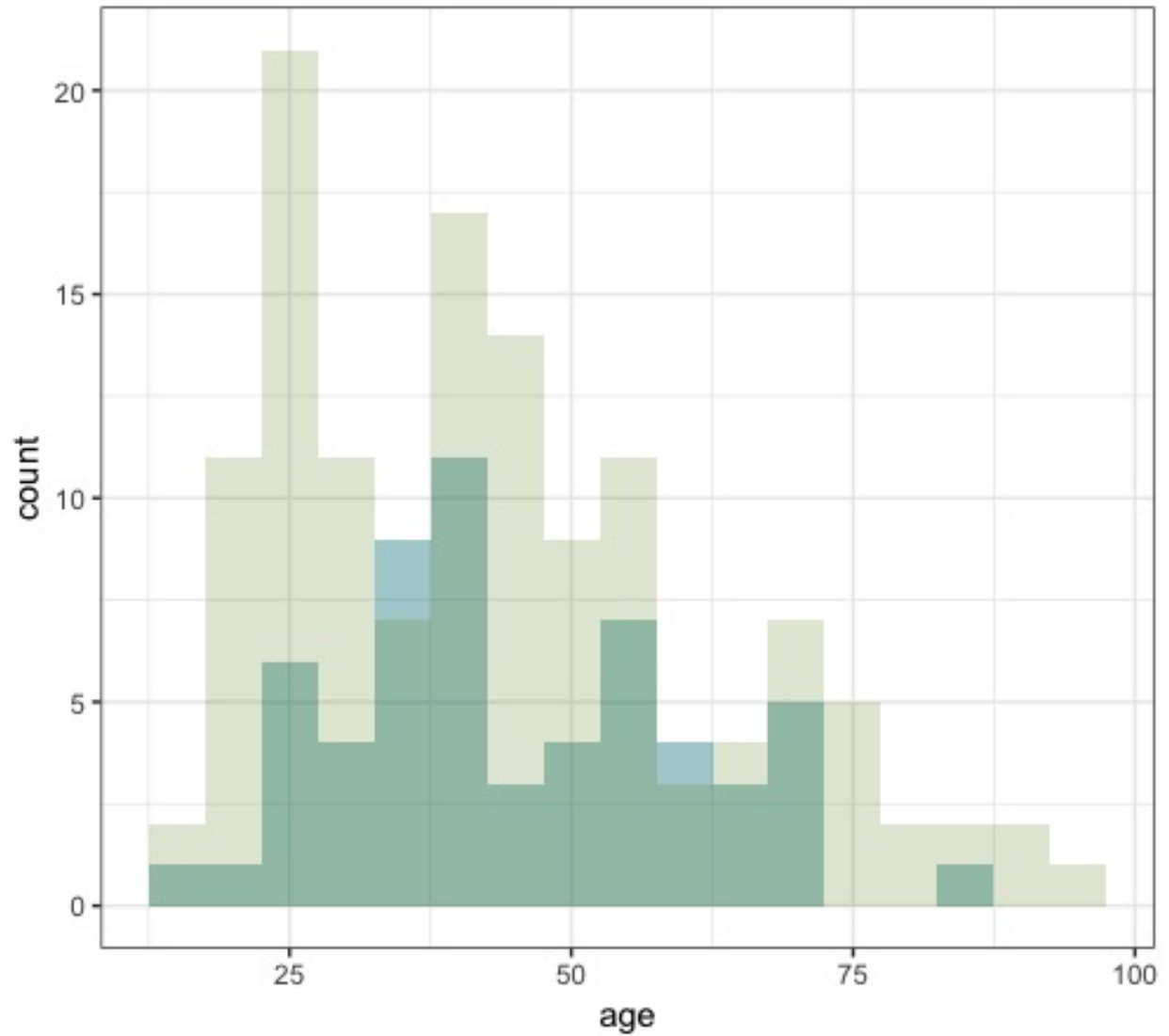
## Supplementary Figures



**Fig. S1.** Density plot of EPP Perception Task. Grey line and shaded region represent mean and standard deviation of responses respectively

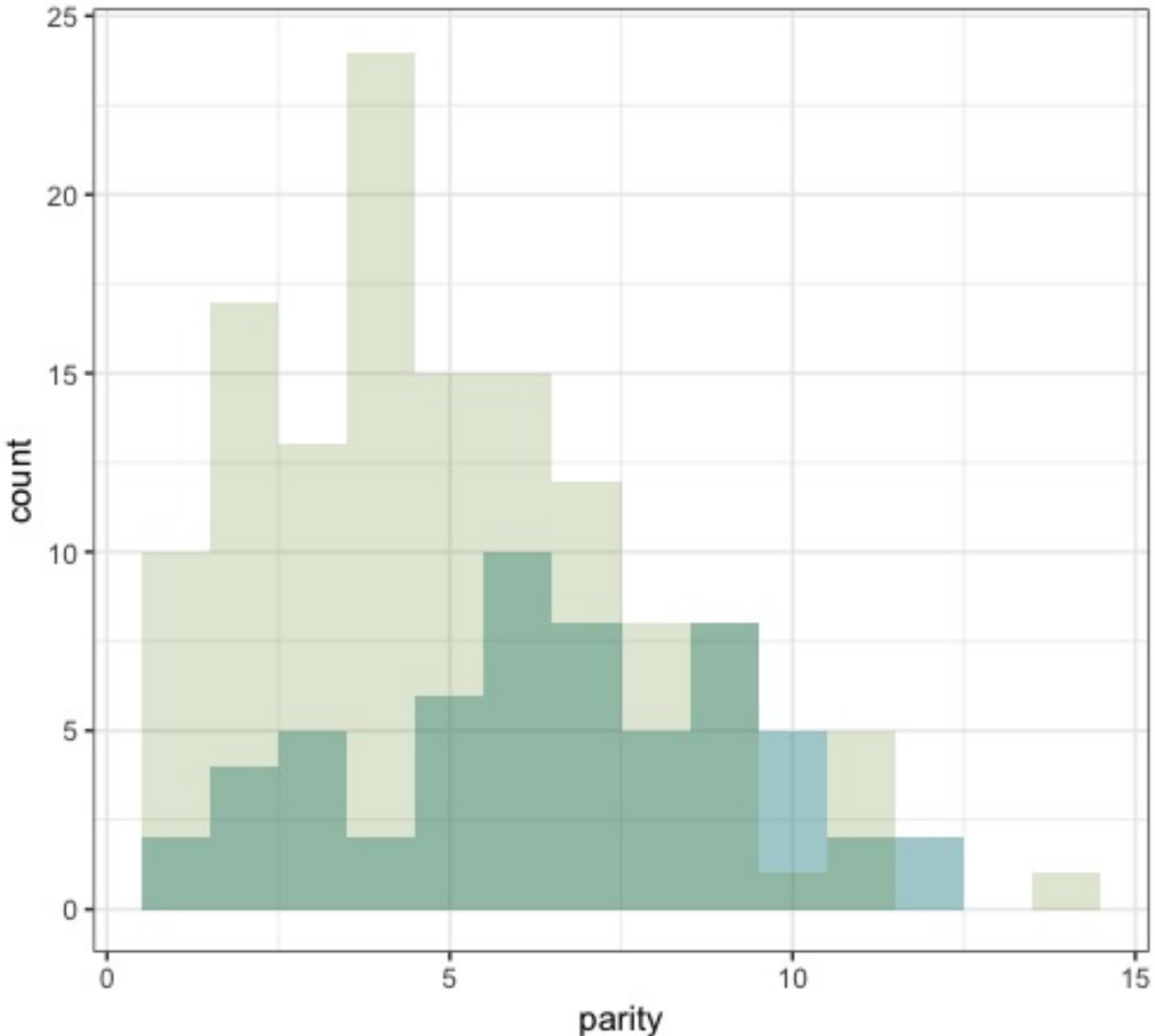


**Fig. S2.** Comparison of age for men in our paternity sample (n=47) with those not living in Omuhonga but not included in our sample (n=68). Comparison population was limited to men who had been married at least once and had at least one child born by a wife during their marriage. Unsampled=light green, Sampled=blue. Overlap shown in dark green.



**Fig. S3.** Comparison of age for women in our paternity sample (n=59) with those not living in Omuhonga but not included in our sample (n=129). Comparison population was limited to women who had been married at least once and had at least one child born during their marriage. Unsampled=light green, Sampled=blue. Overlap shown in dark green.





**Fig. S4.** Comparison of parity for women in our paternity sample (n=59) with those not living in Omuhonga but not included in our sample (n=129). Comparison population was limited to women who had been married at least once and had at least one child born during their marriage. Unsampled=light green, Sampled=blue. Overlap shown in dark green.

## Children per Woman

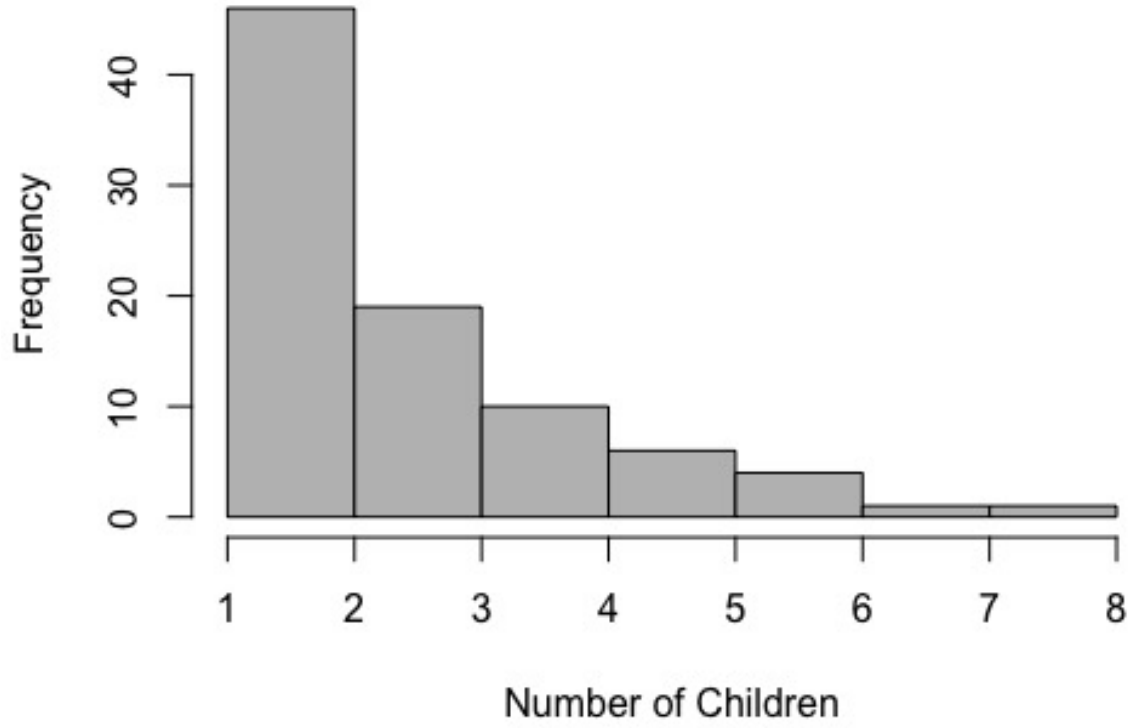


Fig. S5. Frequency distribution showing number of children in the sample, by mother.

## Number of Children Per Social Father

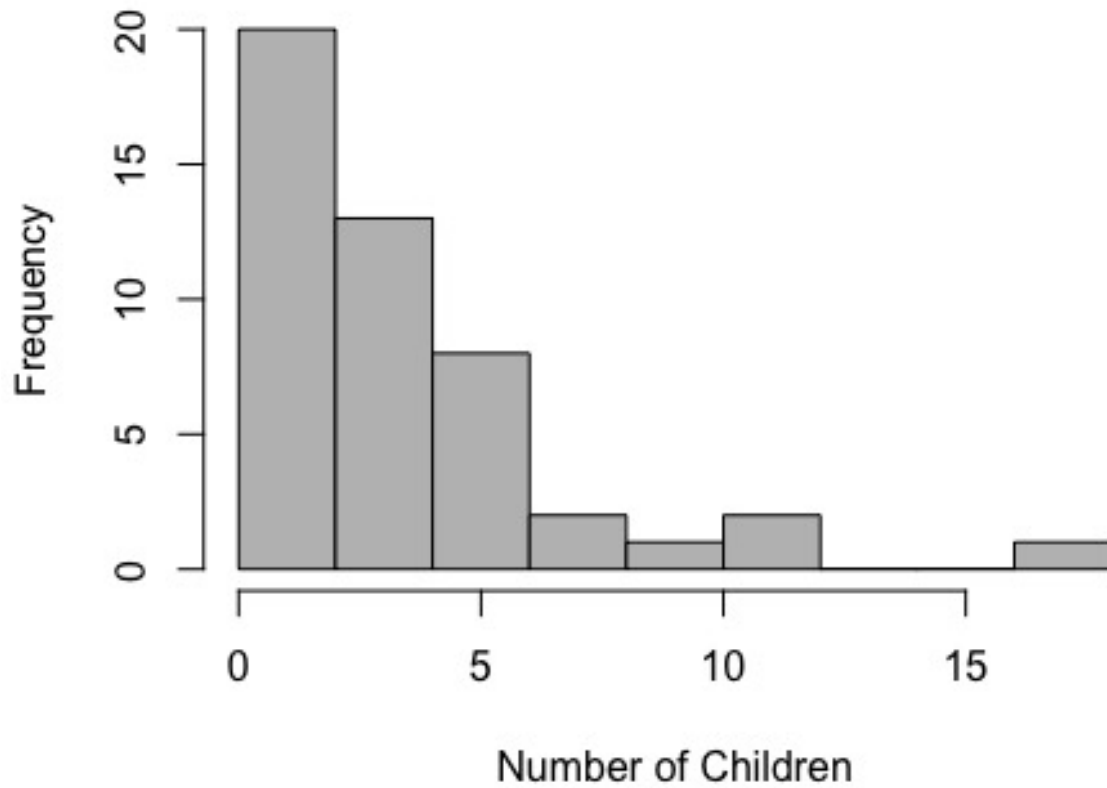


Fig. S6. Frequency distribution showing number of children in the sample, by social father.

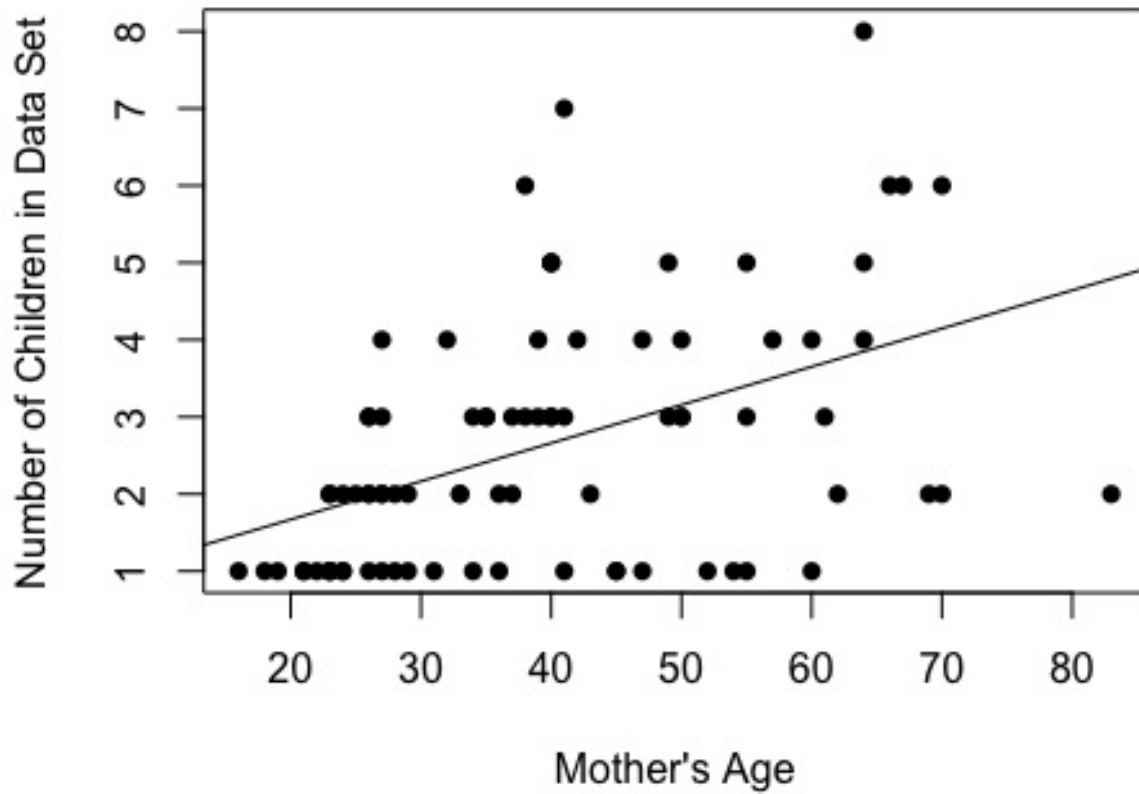
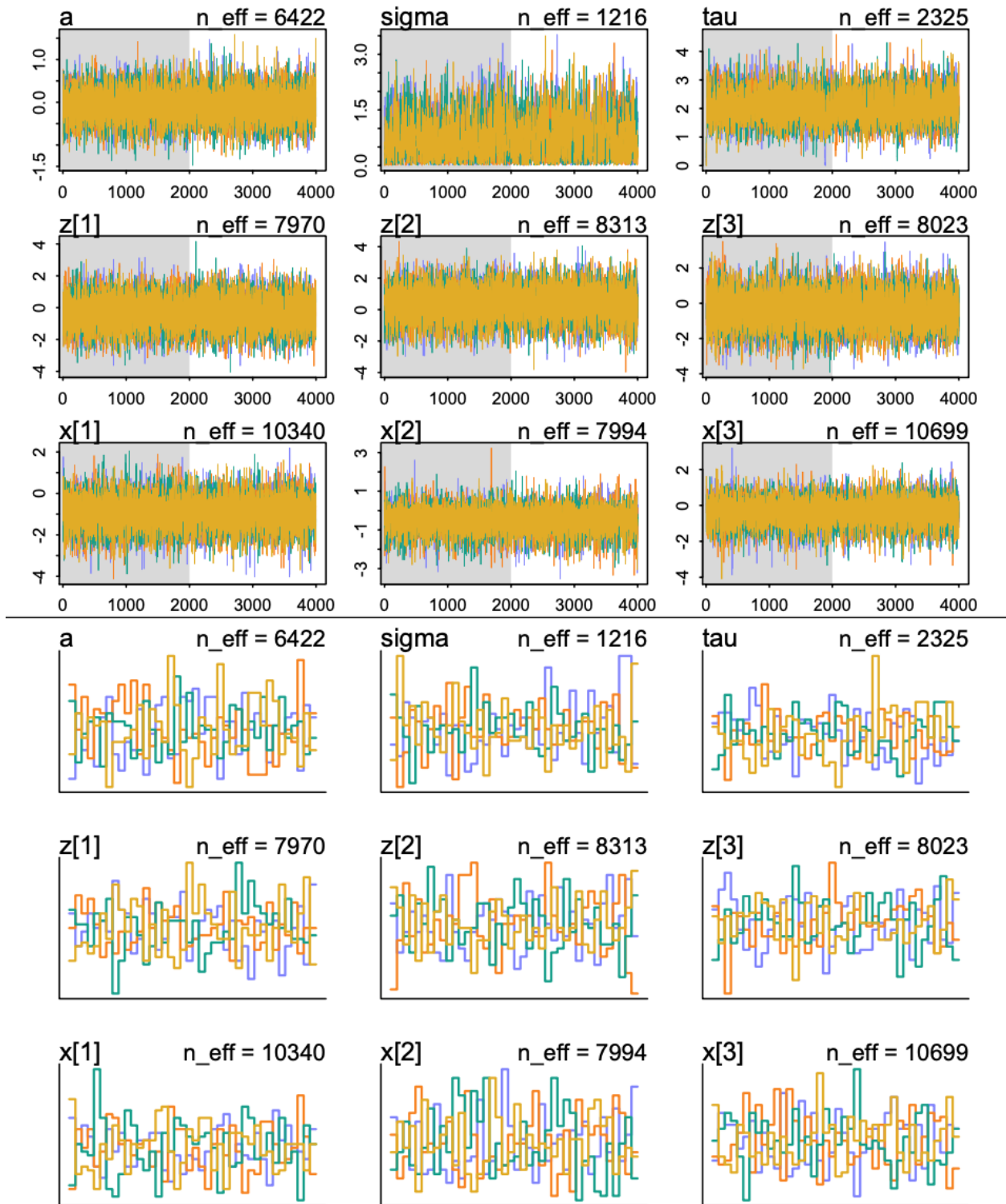
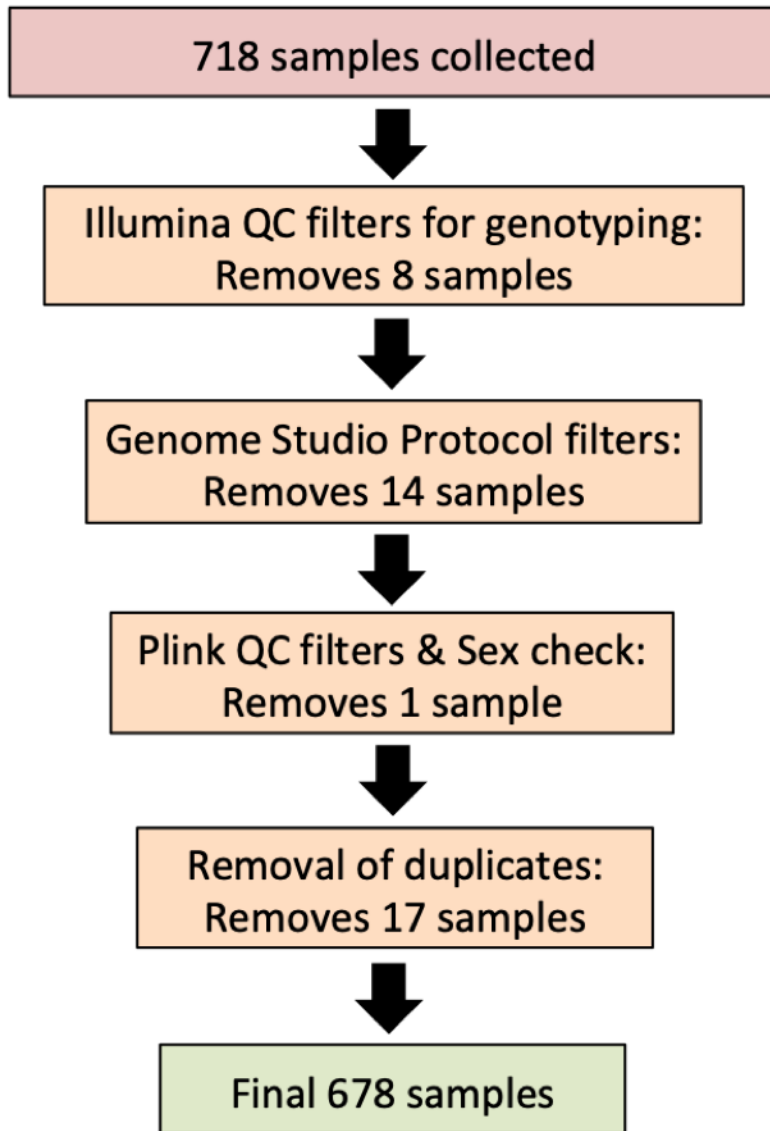


Fig. S7. Number of sampled children, by mother's age.



**Fig. S8.** Trace plots (top) and rank histograms (bottom) for the model that clusters by mother and mother-husband dyad.



**Fig. S9.** Flow chart of DNA sample quality control, beginning with total number of population samples and the final number of samples available for paternity analysis.