Kinship provides the fundamental structure of human society: descent determines the inheritance pattern between generations, whereas residence rules govern the location a couple moves to after they marry. In turn, descent and residence patterns determine other key relationships such as alliance, trade, and marriage partners. Hunter-gatherer kinship patterns are viewed as flexible, whereas agricultural societies are thought to have developed much more stable kinship patterns as they expanded during the Holocene. Among the Bantu farmers of sub-Saharan Africa, the ancestral kinship patterns present at the beginning of the expansion are hotly contested, with some arguing for matrilineal and matrilocal patterns, whereas others maintain that any kind of lineality or sex-biased dispersal only emerged much later. Here, we use Bayesian phylogenetic methods to uncover the history of Bantu kinship patterns and trace the interplay between descent and residence systems. The results suggest a number of switches in both descent and residence patterns as Bantu farming spread, but that the first Bantu populations were patrilocal with patrilineal descent. Across the phylogeny, a change in descent triggered a switch away from patrilocal kinship, whereas a change in residence triggered a switch back from matrilocal kinship. These results challenge “Main Sequence Theory,” which maintains that changes in residence rules precede change in other social structures. We also indicate the trajectory of kinship change, shedding new light on how this fundamental structure of society developed as farming spread across the globe during the Neolithic.

Significance

The agricultural revolution had a dramatic effect on all aspects of human society, but piecing together how humans lived as they spread farming practices worldwide remains difficult. In particular, the fundamental structures of human society, namely the way that property is inherited and the rules governing post-marriage residence, do not leave a clear trace in the archaeological record and, therefore, have been largely intractable. However, the recent availability of phylogenetic language trees coupled with new Bayesian statistical techniques makes it possible to reconstruct the ancestral state of Bantu kinship and reveals that inheritance and residence rules coevolved as farming spread throughout sub-Saharan Africa. Our results question current theory suggesting that residence rules are the primary driver of all other human social structures.
suggestion provides support for “Main Sequence Theory,” which proposes that worldwide there is a pattern of change with residence rules driving change in other social structures (2). In particular, Murdock argues “when any social system undergoes change, such change regularly begins with a modification in the rule of residence” (ref 2, p 221). Changes in descent follow and are always consistent with the change in residence. Changes in kinship terminology are affected by changes in both residence and descent and may follow some considerable time afterward (2).

There is therefore no agreement among researchers about the ancestral states or patterns of change in kinship traits among Bantu societies based on historical linguistic methods. There are a number of problems with relying solely on reconstructions of ancestral vocabulary to infer social organization in the past. First, although linguists use systematic methods of reconstructing proto-forms based on regular sound changes (13, 14), there is still an element of subjectivity in this approach, which can lead different researchers to suggest contradictory results based on the same words (7, 13, 14). Furthermore, the inference of social organization in a past society is a further step removed from this process and assumes a direct relationship between particular words and particular forms of organization. Although it may be possible to reconstruct the sound of a particular lexical item, the meaning of such a word is less clear because of the possibility of semantic shifts (17).

Phylogenetic comparative methods, adapted from evolutionary biology, offer an alternative way of reconstructing the evolutionary history of cultural traits such as kinship structures. These techniques map the traits of interest (in this case descent and residence) onto a phylogenetic tree, which represents the way societies are related to each other historically. The likely forms of these traits in past societies can then be inferred by using an explicit statistical model of trait evolution. Importantly, these reconstructions are probabilistic, meaning it is possible to assess how much confidence to place in any particular reconstruction. The likelihood of alternative hypotheses for the pattern of evolution of traits over the tree can then be estimated. Furthermore, performing analyses over a sample of phylogenetic trees can explicitly incorporate uncertainty about the historical relationships between societies. These methods can be used to estimate the cultural history of a language family, even where historical records or archaeological evidence are absent, and have been used to examine the history of residence patterns in both the Indo-European and Austronesian language families (18, 19).

The Bantu speaking people of sub-Saharan Africa represent one of the major early farming expansions; the advantage of the Indo-European and Austronesian language families (18, 19). Agricultural food production allowed a single language group to displace and expand into the lands of previous hunter-gatherer populations (6). This process has allowed for language phylogenies of Bantu societies to be generated (9, 20–22), which can be combined with comprehensive data on extant kinship patterns across a large number of Bantu cultures (23, 24) to infer historical patterns of cultural evolutionary change. A previous attempt to infer the ancestral state of descent for the original Bantu population (25, 26) was inconclusive, possibly because of the small sample size, the single phylogeny (20), or the maximum likelihood methods used.

Here, we advance previous attempts to reconstruct the kinship traits across the Bantu language family by applying Bayesian phylogenetic comparative methods (27–29) to a dataset of 122 Bantu ethnolinguistic groups to infer the ancestral state and evolutionary trajectory of Bantu kinship patterns. Analysis of basic vocabulary items (30) has enabled the phylogenetic relationships among more than 500 Bantu languages to be inferred (9), which has increased the scope for the number of different Bantu cultures that can be analyzed and enabled the incorporation of groups that were underrepresented in previous analyses. In particular, the current sample includes a number of additional societies from the “Northwest” and “Forest West” regions, which may be particularly important in inferring trait states at the earliest nodes in the trees. Furthermore, Bayesian phylogenetic methods, which can incorporate uncertainty about the phylogenetic relationships between cultures, provide a more accurate picture of the inferences that can be drawn from comparative data. In this way, it is possible to infer the ancestral states of kinship traits and, hence, reconstruct the cultural history of the Bantu expansion, which means it is possible to estimate the likely order of trait change and test whether changes in residence rules precede changes in inheritance patterns as proposed by Main Sequence Theory (2, 7, 16).

Results

We found both descent and residence patterns showed strong phylogenetic signal among Bantu-speaking cultures. A phylo.d test (31) revealed that both traits have phylogenetic patterning that is extremely unlikely to occur by chance (P < 0.001) and are not significantly different from Brownian motion (SI Appendix, Table S2), which supports models in which descent and residence patterns each evolve along the branches of the Bantu language tree.

We inferred the ancestral states for kinship traits at the root of the Bantu language tree by using a reversible jump (RJ) Markov chain Monte Carlo (MCMC) method, which simultaneously explores the posterior sample of phylogenetic trees and possible models of evolution of traits on the tree, visiting models in proportion to their posterior probability (32). This method allows for uncertainty in the phylogeny and the stochastic nature of the evolutionary process. We find strong support (P = 0.995) for patrilocal descent at the root of the phylogeny, whereas for residence, there was a high probability (P = 0.999) that the root was patrilocal (SI Appendix, Table S3).

A previous study (25) was unable to resolve the ancestral state for descent, but showed that the adoption of cattle herding led to switches to patrilineal root. To check for this false result, we performed a correlated evolution analysis of descent and cattle herding in the Discrete package in BayesTraits (27). Both the independent analysis, where the traits are forced to evolve separately, and the dependent analysis, where they can evolve together, showed strong support for patrilineal root without cattle herding as the ancestral Bantu state, indicating that the later adoption of cattle did not falsely influence the root reconstruction for descent.

The reconstruction (Fig. 1) of the probability of states at ancestral nodes across the phylogeny suggests that from patriliney as the descent pattern at the root (A) of the Bantu phylogeny, there was an initial switch to matriliney (B) in a few Forest West Bantu cultures (following ref. 35) and those in the West Savannah. As Bantu populations spread further across into East Africa and then South, there was a subsequent shift back to patriliney (D), followed by a switch to matriliney (E) in the South East Bantu and then a shift to mixed descent (F) and a final change back to patriliney (G) as Bantu populations reached the southern tip of Africa (Fig. 1). Residence followed a similar pattern to descent, except that there was a secondary switch back to patrilocal at (C) instead of (D), and the final switch to patrilocal came at (F) not (G) (SI Appendix, Fig. S4). Neocolony is shown to be a recent phenomenon supporting earlier research (36). The pattern for both descent and residence is so similar as to suggest correlated evolution between these two traits as Bantu-speaking populations spread across sub-Saharan Africa.

To test for correlated evolution between descent and residence and assess whether changes in residence were driving changes in descent, as proposed by Main Sequence Theory (2), the RJ procedure in the Discrete package in BayesTraits (27) was
used. A log_{10} Bayes factor of 3.11 suggests decisive support for correlated evolution between descent and residence across the Bantu phylogeny (SI Appendix, Table S4).

Fig. 2 shows the reversible-jump analysis of correlated evolution between the traits. There are two routes from the ancestral state of patriliny and patrilocality to the derived state of matri/bilateral descent and matri/neolocal residence: the one in which descent changes first (transitions q_{13} and q_{34} via state C) and the one where residence changes first (q_{12} and q_{24} via state B) (SI Appendix, Fig. S9). In the reversible jump analysis, the route where descent changes first (transitions q_{13} and q_{34}, via intermediate state C) is more likely, with both transitions assigned to zero in none or a very small proportion of the posterior distribution. This result suggests that, for changes from the ancestral state to the derived state, descent changed first, and therefore, because the traits show correlated evolution, probably drove changes in residence. However, there were a number of switches between states in both descent and residence, and for the transitions in the other direction (i.e., from matriliny/matrilocality to patriliny/patrilocality), residence changed first. To check these findings, likelihood comparisons were made between models where the transition rates of the two routes between patrifocal and matrifocal cultures (via state B or state C in Fig. 2) were each forced to zero. The transitions via state C were shown to be decisively more likely, providing strong evidence to support the reversible-jump analysis (SI Appendix, Tables S5 and S6), which suggests that changes in descent drove shifts away from patrifocal cultures, whereas residence drove reversals from matrifocal back to patrifocal cultures. In other words, in moves between patrifocal (A) and matrifocal (D) cultural states, an intermediate state of matriliny with patrilocality (C) is much more likely than an intermediate state of patriliny with matrilocality.

Fig. 1. Ancestral reconstruction of descent across Bantu cultures. The pies at each node show the mean probability of each state from the MCMC analysis plotted on the consensus tree. Gray in the pies shows uncertainty that the node exists in the tree sample. Letters (A–G) denote switch points in the state of descent along the expansion route on the phylogeny, for comparison with the residence trait (SI Appendix, Fig. S4). Cultural groups follow Guthrie (32).
emerge because of the loss of cattle, tsetse fly infestation (26), or some other cause (40), which may have reduced the pressure for inheritance through the male line. The results presented here suggest that, in turn, changes in descent may have driven residence patterns to change to matrilocal.

Overall the phylogenetic comparative approach taken here to infer social organization in past societies offers improvement over relying completely on historical linguistic reconstructions. In particular, this approach allows direct inference of the ancestral states of the particular traits of interest, rather than tackling the issue obliquely via kinship terminology. It would be interesting to reconstruct kinship terminology explicitly using phylogenetic comparative methods to examine the relationship with direct estimates of social organization and as a point of comparison with traditional methods. Archaeological information could also potentially be of use in inferring kinship structures in past human societies. Although many aspects of social organization do not leave direct material remains in the archaeological record, it may be possible to make plausible inferences based on, for example, mortuary practices (41, 42). Unfortunately, the current archaeological record in Africa is incomplete because of the environmental conditions in the equatorial rainforest, which are particularly problematic for the preservation of the remains of human societies of the past (35).

However, the results reported here do at least provide predictions about the likely forms of social organization in past societies that could in theory be tested as archaeological data accumulate.

Our results clearly demonstrate that both residence and descent, which are the key aspects of kinship, coevolved as Bantu populations moved through sub-Saharan Africa. As well as identifying correlated evolution between these traits, the model of evolution indicates that descent changed first for moves away from patrilocal states, whereas residence changed first for moves back from matrilocal states. When inheritance through the male line breaks down, so that wealth is no longer passed from father to son, women may be more likely to stay with their relatives, postmarriage, to garner support from close relatives in tasks such as child rearing. Indeed, communal breeding by related females has been shown to promote matrilineality (43). That descent could change before residence questions Main Sequence Theory, which proposes that it is changes in residence rules that drive changes in the whole social structure of a society (2, 16). In the Bantu at least, our results suggest that residence only changes away from a matrifocal pattern, which may indicate that the costs and benefits of changing kinship patterns differ depending on whether the culture is patrifocal or matrifocal.

Methods
Data. Ethnographic data were derived from the Ethnographic Atlas (23, 24) and matched to the language data by name and geographic location (following ref. 19).

For the descent pattern variable 43 of the Ethnographic Atlas (23, 24), descent: major type, was used. Among Bantu-speaking populations the six states featured are as follows: 1, patrilineal; 2, dualateral; 3, matrilineal; 4, quasi-lineages; 5, amblilineal; 6, mixed (SI Appendix, Table S1). For ancestral state analyses, the three states used were as follows: patrilineal, matrilineal, and bilateral (including states 2, 4, 5, and 6 above) (following ref. 25). SI Appendix, Fig. S1 shows the descent pattern across extant Bantu cultures. A binary variable was used in the coevolution analyses: patrilineal (ancestral state) versus matrilineal or bilateral, i.e., patrilineal vs. any other state (derived states from Multistate analysis).

For the residence pattern variable 11 of the Ethnographic Atlas (23, 24), transfer of residence at marriage: after first years, was used (following refs. 18 and 19). This variable includes avunculocality within “husband to wife’s group,” but excludes a short stay after marriage by the husband in the wife’s group to perform bride service, capturing the permanent move of one or both marriage partners to a new location (and enables comparison with phylogenetic studies of residence patterns in other language families; refs. 18 and 19). Among Bantu speaking populations, the following three states

Discussion
Our results show that even when archaeological and historical evidence are lacking, it is possible to go beyond historical linguistic inference, which is highly contested in the Bantu case (7, 12–14), and reconstruct the ancestral states of kinship traits throughout the Bantu expansion. Our analyses, based on the most comprehensive Bantu language phylogeny available (9), suggest that the ancestral Bantu society, which was located in the Benue Valley on what is now the Nigerian/Cameroon border, was both patrilineal and patrilocal. As far as residence is concerned, this result is consistent with genetic studies on sub-Saharan farmers that are argued to imply a patrilocal history; fast evolving polymorphisms indicate that the ratio of mtDNA to Y-chromosome variation is much higher in farming than hunter-gatherer populations, suggesting females moved residence before they reproduced, whereas males stayed in their natal locality (37–39).

The results of this study do not support the conclusions of either of the historical linguistic studies of the Bantu (7, 12–14). Instead of the bilateral and ambilocal kinship proposed by Vansina (12), or the matrilineal and matrilocal kinship proposed by Marck and coworkers (7, 14), our analyses support yet another hypothesis, providing evidence for the first time to our knowledge that the kinship pattern for the ancestral Bantu population was both patrilineal and patrilocal. Furthermore, the phylogenetic methods used enable us to postulate that for both traits, there were switches back and forth between patri- and matri-focal and matri-focal states as the Bantu populations expanded throughout sub-Saharan Africa. These switches suggest that there was some flexibility in these traits as Bantu populations spread southwest, but that unilineal descent systems were retained throughout Bantu history, in contrast to the scenario proposed by Vansina (13).

Despite being unable to resolve the ancestral state of descent among the Bantu, Holden and Mace (25, 26) were able to suggest that patriliney is associated with pastoralism. Matriliney could
are featured 1, wife to husband’s group (patrilocal); 2, couple to either group or neolocal; 3, husband to wife’s group (matrilocal) (SI Appendix, Table S1). The data set is available in Zenodo. 50 shows the MCC tree, with a clade credibility pattern across all Bantu cultures. A binary variable was used in the coevolution analyses: patrilocal (ancestral state) versus matrilocal or bineolocal i.e., patrilocal vs. all other states (derived states from Multistate analysis).

**Phylogeny.** We used the most comprehensive Bantu phylogeny available to date based on linguistic data from 542 Bantu languages (9), and consisting of 500 phylogenies from a Bayesian posterior distribution of trees. A maximum clade credibility tree was derived from the full sample, for illustrative purposes, using TreeAnnotator (44) and displayed by using FigTree (44). A previous analysis (45) using a different Bayesian Bantu phylogeny, based on a different dataset, found similar results to those reported here.

**Analyses.** Phylogenetic signal in data measures the extent to which related taxa are more similar in a particular trait than would be expected by chance. If the signal were low, then it would be difficult to infer the evolutionary history of that trait because the phylogeny gives little or no information on the distribution of the trait among extant populations. The signal in each trait was tested by calculating D using the function phylo.d (31) in the Caper package (46) in R (47). A D value significantly different from random indicates that related populations are more similar in a particular trait than would be expected by chance (31).

A likelihood framework together with Bayesian inference was used to model the evolution of traits along the branches of the tree. Analyses were carried out in BayesTraits (available from www.evolution.rdg.ac.uk) (27, 32) using the RJ MCMC procedure. Maximum likelihood analyses of the data, which give point estimates of model parameters, indicated that the prior distributions (the prior belief about the model of evolution) could be best described by an exponential probability distribution because of the small number of changes per unit of branch length. The priors were seeded from exponential hyperpriors (a distribution from which the prior is drawn) with a mean and variance in the range of 0–10 (32). Convergence was checked visually by evaluating changes in the log-likelihood in nodes from the run with the median likelihood taken from the post-transition models, rate parameters, log likelihoods, and states at ancestral that convergence had been reached. The posterior probabilities for the MCMC chain was run five times for 5 million iterations sampled every 100, carried out in model the evolution of traits along the branches of the tree. Analyses were

**Discrete, a procedure in BayesTraits (27, 32), was used to test hypotheses about correlated evolution between two binary traits. Two models are fitted in this procedure: an independent model in which the two traits evolve over the phylogeny, independent of the state of the other trait, and a dependent model in which the state of one trait affects the probability of change in the other trait. A Bayes Factor (BF) (50) comparison is made between the independent and the dependent models such that independent evolution can be rejected if there is support for the dependent model; support for the dependent model indicates correlated evolution between the traits. For comparisons between models, the log$_{10}$ BF was calculated in Tracer (48), as twice the difference in the marginal likelihood (32). The BF shows the weight of evidence to support one model over another, from 0–0.5 (in-substantial), to 0.5–1.0 (substantial), to 1.0–2.0 (strong), to >2.0 (decisive) (50). Comparing the transition rates between states indicates the relative timing of the evolution of traits. This method allows for an evaluation of directionality such that the order of evolution can be inferred; by showing which trait evolved first it is possible to test hypotheses explicitly.

**ACKNOWLEDGMENTS.** C.O., T.C., and R.M. are supported by European Research Council Advanced Grant (AdG 249347). C.O. is also supported by a Leverhulme Early Career Fellowship. S.S. is supported by a Royal Society University Research Fellowship. Q.D.A. is supported by a Rutherford Discovery Fellowship.

---


---

To find the posterior probability distributions of states at ancestral nodes across the phylogenies, an RJ MCMC analysis was run with the additional use of the addNode command in BayesTraits, which reports the posterior distribution of trait states at each internal node on the phylogenetic tree (32). The ancestral state reported for each node of the tree is the combined posterior probability of each state at that node with the posterior probability that the node itself exists in the tree and is plotted on the consensus tree.

The RJ MCMC method used simultaneously travels through the full posterior sample of trees and the universe of all possible models of evolution visiting those models in proportion to their posterior probability. This method provides a posterior probability sample of models of evolution and the rate parameters of the models themselves (27, 49) by visiting the models in proportion to their likelihood and assesses which transition rates, between states, are set to zero (Z) and which are equal to other rate parameters (0, 1, 2).
Supplementary Information Appendix

The terms used in this paper follow standard definitions (1, 2), see Table S1.

Table S1 Kinship Terminology

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Descent</em></td>
<td></td>
</tr>
<tr>
<td>Patriliny</td>
<td>Descent traced through the male line</td>
</tr>
<tr>
<td>Matriliny</td>
<td>Descent traced through the female line</td>
</tr>
<tr>
<td>Unilineal</td>
<td>Descent traced through either male or female line</td>
</tr>
<tr>
<td>Bilateral</td>
<td>Descent traced equally through the both male and female line</td>
</tr>
<tr>
<td>Duo lateral</td>
<td>Descent traced through the male or female line, situation dependent</td>
</tr>
<tr>
<td>Quasi-lineages</td>
<td>Filiation based, not descent</td>
</tr>
<tr>
<td>Ambilineal</td>
<td>Choice over descent</td>
</tr>
<tr>
<td>Mixed descent</td>
<td>No clear descent pattern</td>
</tr>
<tr>
<td><em>Post-marital Residence</em></td>
<td></td>
</tr>
<tr>
<td>Patrilocal</td>
<td>Couple reside with husband’s kin</td>
</tr>
<tr>
<td>Matrilocal</td>
<td>Couple reside with wife’s kin</td>
</tr>
<tr>
<td>Neolocal</td>
<td>Couple reside in new area, not with kin</td>
</tr>
<tr>
<td>Bilocal</td>
<td>Couple reside with either kin</td>
</tr>
</tbody>
</table>
Bantu Cultures

The descent and residence pattern for all extant Bantu-speaking cultures in the sample are shown in Figure S1 and Figure S2 respectively.

Figure S1 Distribution of extant Bantu-speaking cultures by mode of descent
Figure S2 Distribution of extant Bantu-speaking cultures by mode of residence

Phylogeny

We used the most comprehensive Bantu phylogeny available to date, based on linguistic data from 542 Bantu languages (3), and consisting of 500 phylogenies from a Bayesian posterior distribution of trees.
**Phylogenetic signal**

Phylogenetic signal was tested in the descent and residence traits (Table S2) by calculating D using the function `phylo.d` (4) in the Caper package (5) in R (6). A D value significantly different from random indicates that related populations are more similar in a particular trait than would be expected by chance (4).

**Table S2 D statistic for phylogenetic structure of binary traits.**

<table>
<thead>
<tr>
<th>Trait Data</th>
<th>Est. D</th>
<th>Probability of Est. D resulting from:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No phylogenetic structure (random)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brownian motion</td>
</tr>
<tr>
<td>Descent</td>
<td>0.22</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.214</td>
</tr>
<tr>
<td>Residence</td>
<td>0.10</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.382</td>
</tr>
</tbody>
</table>

**Ancestral States**

The state of each trait at the root of the phylogeny was tested using Multistate in BayesTraits (7, 8). The results are shown in Table S3.

**Table S3 Mean probability of ancestral state of kinship traits at the root of the Bantu phylogeny.**

<table>
<thead>
<tr>
<th>Trait</th>
<th>Log Likelihood</th>
<th>Ancestral State</th>
<th>Mean probability</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descent</td>
<td>-96.93</td>
<td>Patriliny</td>
<td>0.995</td>
<td>0.00004</td>
</tr>
<tr>
<td>Residence</td>
<td>-76.43</td>
<td>Patrilocality</td>
<td>0.999</td>
<td>0.00005</td>
</tr>
</tbody>
</table>

Figures S4 and S6 show the ancestral nodes for residence reconstructed on a phylogeny and geographic locations, respectively, for comparison with the descent traits shown in Figures 1, S3 and S5.
Figure S5 and S6 show, for illustrative purposes, the mean probability of each state at each ancestral node of the Bantu phylogeny for descent and residence respectively, plotted at locations on the map of sub-Saharan Africa where the ancestral culture has been inferred (3). The route (shown in black) that the expansion-wave of Bantu-speaking populations took from their ancestral home on the current Nigeria/Cameroon border (A) first towards East Africa, via the western rain forest (B), and then down to the southern tip of Africa (G) is shown (from 3). The points at which each trait switches between states are shown (A-G) as well as other selected points along the expansion route.
Figure S3. Ancestral reconstruction of descent across Bantu cultures. The pies at each node show the mean probability of each state from the MCMC analysis plotted on the consensus tree. Grey in the pies shows uncertainty that the node exists in the tree sample. Letters (A-G) denote switch points in the state of descent along the expansion route on the phylogeny, for comparison with the residence trait (Figure S4). Cultural groups follow Guthrie (32).
Figure S4 Ancestral reconstruction of residence across Bantu cultures. The pies at each node show the mean probability of each state from the MCMC analysis plotted on the consensus tree. Grey in the pies shows uncertainty that the node exists in the tree sample. Letters (A-G) denote switch points in the state of residence along the expansion route on the phylogeny, for comparison with the switches in the descent trait (Figure 1 & S3). Cultural groups follow Guthrie (9).
Figure S5 Ancestral reconstruction of descent across Bantu cultures with geographical location and expansion-wave route shown. (Grey in pie shows uncertainty that the node exists from the tree sample). Letters (A-G) denote switch points in the state of descent along the expansion route on the phylogeny, for comparison with the phylogenetic position of switches (Figure 1 & S3) and the switches in the residence trait (Figures S4 & S6).
Figure S6 Ancestral reconstruction of residence across Bantu cultures with geographical location and expansion-wave route shown. (Grey in pie shows uncertainty that the node exists from the tree sample). Letters (A-G) denote switch points in the state of residence along the expansion route on the phylogeny, for comparison with the phylogenetic position of switches (Figure S4) and the switches in the descent trait (Figures 1, S3 & S5).
**Evolutionary Models**

The reversible-jump procedure in *BayesTraits* visits the universe of all possible models in proportion to their posterior likelihood (8). Figure S7 and Figure S8 show the highest ranked model across the posterior and transition rates between states across the whole posterior distribution.

The highest-ranked model of descent in the posterior probability distribution (accounting for 46% of the distribution) shows that from bi-lateral descent there are no transitions to matrilineal descent (Figure S7). Across the full posterior probability distribution this rate is zero (Z) 56% of the time, while all other transition rates, including direct shifts from patriliney to matriliney and back, are strong and never assigned to zero (Z).

For residence, there are fast rates out of the ancestral state of patrilocality to both other states and back to patrilocality directly from matrilocality (Figure S8). The transition rates out of neo-locality and from matrilocality to neo-locality are weaker. Residence is also a flexible trait across the Bantu phylogeny, but the strongest rates are between patrilocal and matrilocal residence and from patrilocality to neo-locality.
Figure S7 Estimated rate parameters of evolution of descent among Bantu populations under the RJMCMC analysis. Z denotes visits assigned to zero as a proportion of posterior probability distribution. Thickness of arrows reflects proportion of time transition rate is not assigned to zero. Dashed line denotes a zero rate in the top-rated RJ-derived model. Number below rate name ($q_{ij}$) is the mean transition rate where rate distribution approximates normal.
Figure S8 Model of evolution for residence among Bantu populations. Z denotes visits assigned to zero as a proportion of posterior probability distribution. Thickness of arrows reflects proportion of time transition rate is not assigned to zero. Number below rate name ($q_{ij}$) is the mean transition rate where rate distribution approximates normal.

**Correlated Evolution**

*Discrete*, a procedure in *BayesTraits* (7, 8), was used to test for correlated evolution between two binary traits. A $\log_{10}$ Bayes Factor (BF) (10) comparison can be made between the independent and the dependent models such that independent evolution can be rejected if there is support for the dependent model; support for the dependent model indicates correlated evolution between the traits.

We were concerned that the adoption of cattle may have influenced the result of the ancestral state analysis at the root for descent. The Multistate analysis for descent shows strong probability of patriliny at the root (0.995), while for cattle, there is a strong probability of no cattle at the root (0.997). A Discrete analysis of the
two traits together shows a high probability of patriliny (0.991) and no cattle (0.995) at the root in the independent model, where the traits are constrained to evolve separately, but also in the dependent model (probability of patriliny/no cattle 0.987), where they are allowed to evolve together. This suggests that although there is substantial support (log10 Bayes factor of 0.9) for correlated evolution between these traits, supporting a previous study (11), the adoption of cattle has not influenced the reconstruction of patriliny as the ancestral state for descent.

The Discrete analysis of descent and residence shows decisive support for correlated evolution between these two traits (Table S4).

<table>
<thead>
<tr>
<th>Correlated evolution</th>
<th>Dependent model</th>
<th>Independent model</th>
<th>Log10 Bayes Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Log likelihood</td>
<td>S.E.</td>
<td>Log likelihood</td>
</tr>
<tr>
<td>Descent - Residence</td>
<td>-121.45</td>
<td>+/- 0.04</td>
<td>-128.62</td>
</tr>
</tbody>
</table>

**Timing of Trait evolution**

Transition rates between states of two binary traits can indicate the relative timing of trait evolution and therefore make inferences about causality. In Figure S9 there are two routes from the ancestral state [0,0] to the derived state [1,1] for both characters, via [0,1] (transitions $q_{12}$ and $q_{24}$) and via [1,0] (transitions $q_{13}$ and $q_{34}$). In order to test which of the traits evolved before the other, a comparison is made between the two routes (8). If the route via [0,1] (transitions $q_{12}$ and $q_{24}$) is stronger than the route via [1,0] (transitions $q_{13}$ and $q_{34}$) this indicates support for a change in trait two preceding a change in trait one. If there is correlated evolution between
the two traits this indicates that the change in trait two is implicated in the evolution of trait one.

![Diagram of correlated evolution between two binary traits](image)

**Figure S9** Correlated evolution between two binary traits. There are two routes from the ancestral state (0,0) to the derived state (1,1), either via (0,1) a change in state two first, or (1,0) a change in state one first (it is assumed that the states do not change at the same time over a sufficiently small time period) (8).

To test the findings drawn from Figure 2, we compared likelihoods between models where the transition rates of each of the two routes from the ancestral (A) to the derived (D) states were forced to zero. The model with route via intermediate state C forced to zero had a decisively lower likelihood than the route via intermediate state B, suggesting that C is the important route, since forcing it to zero disrupts the model much more (Table S5).
For transitions in the other direction (matrifocal to patrifocal) we compared likelihoods between models where the transition rates of each of the two routes from the derived (D) to the ancestral (A) states were forced to zero (Figure 2). The model with route via intermediate state C forced to zero had a decisively lower likelihood than the route via intermediate state B, suggesting that C is the important route (Table S6).

### Table S6 Comparison of transition rates forced to zero – Matrifocal to Patrifocal.

<table>
<thead>
<tr>
<th>Model</th>
<th>Log Likelihood</th>
<th>S.E.</th>
<th>Log_{10} Bayes Factor with Model C=0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transition rates via state C are zero</td>
<td>-132.79</td>
<td>0.241</td>
<td>-</td>
</tr>
<tr>
<td>Transition rates via state B are zero</td>
<td>-122.60</td>
<td>0.109</td>
<td>4.43</td>
</tr>
</tbody>
</table>
Supplementary References