Supplementary Information Appendix

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

Table S1 Kinship Terminology

Term		Definition
Descent	Patriliny	Descent traced through the male line
	Matriliny	Descent traced through the female line
	Unilineal	Descent traced through either male or female line
	Bilateral	Descent traced equally through the both male and female line
	Duo lateral	Descent traced through the male or female line, situation dependent
	Quasi-lineages	Filiation based, not descent
	Ambilineal	Choice over descent
	Mixed descent	No clear descent pattern
Post-marital Residence	Patrilocal	Couple reside with husband's kin
	Matrilocal	Couple reside with wife's kin
	Neolocal	Couple reside in new area, not with kin
	Bilocal	Couple reside with either kin

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.

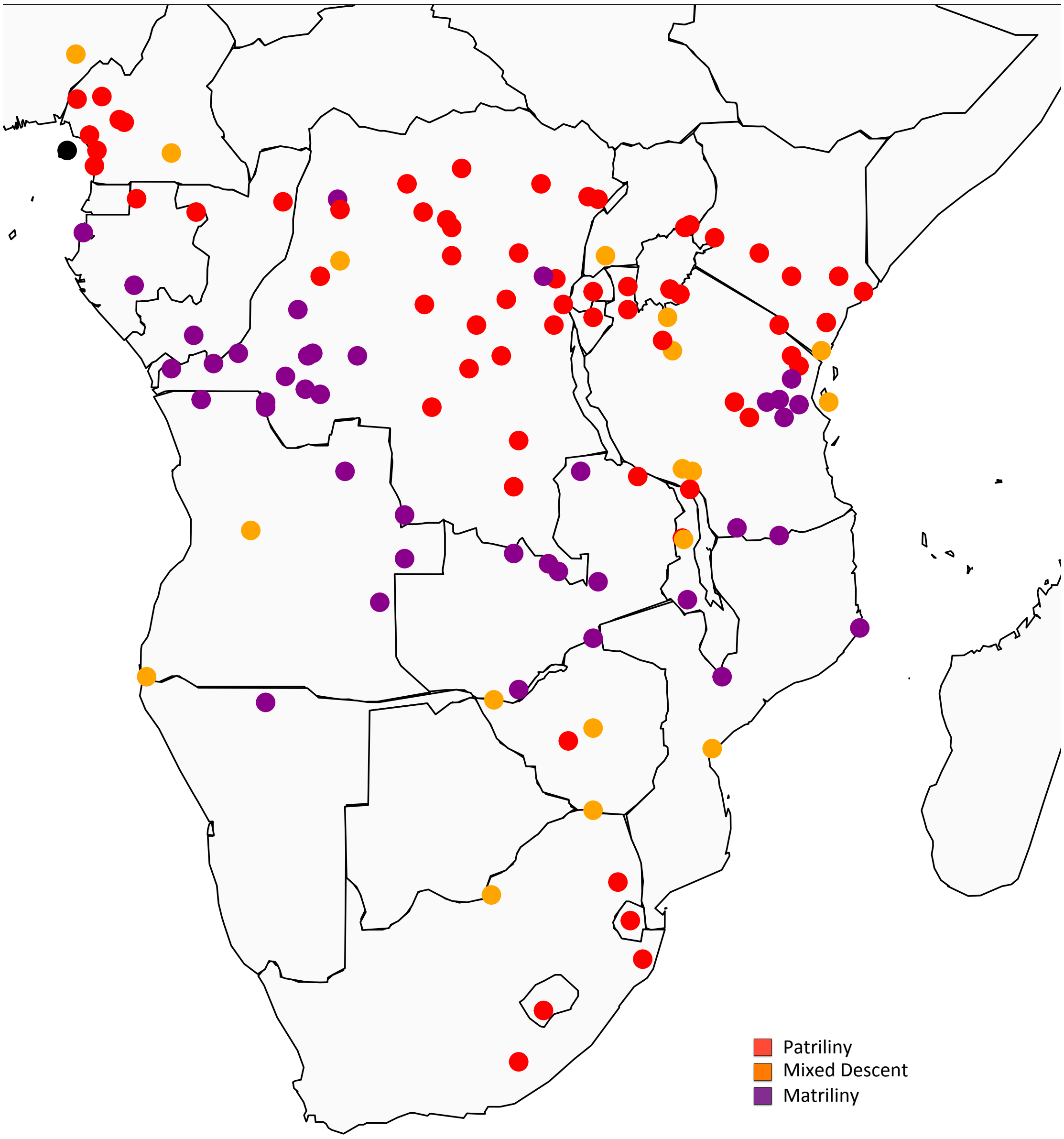

Figure S1 Distribution of extant Bantu-speaking cultures by mode of descent

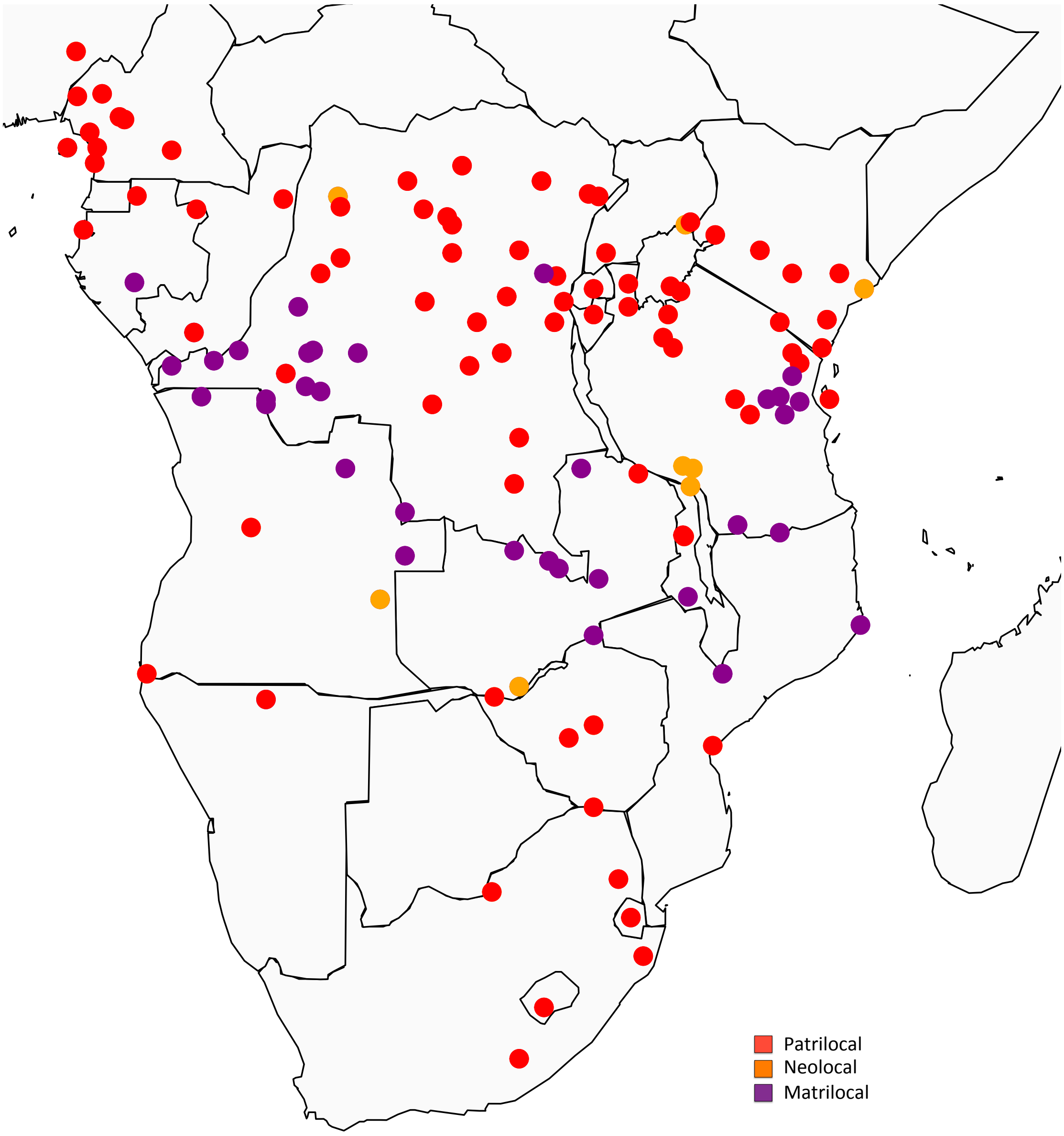

Figure S2 Distribution of extant Bantu-speaking cultures by mode of residence

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.

Trait Data	**Est. D**	**Probability of Est. D resulting from:**	
		No phylogenetic structure (random)	Brownian motion
Descent	0.22	0.000	0.214
Residence	0.10	0.000	0.382

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.

Trait	Log Likelihood	Ancestral State	Mean probability	S.E.
Descent	-96.93	Patriliny	0.995	0.00004
Residence	-76.43	Patrilocality	0.999	0.00005

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.

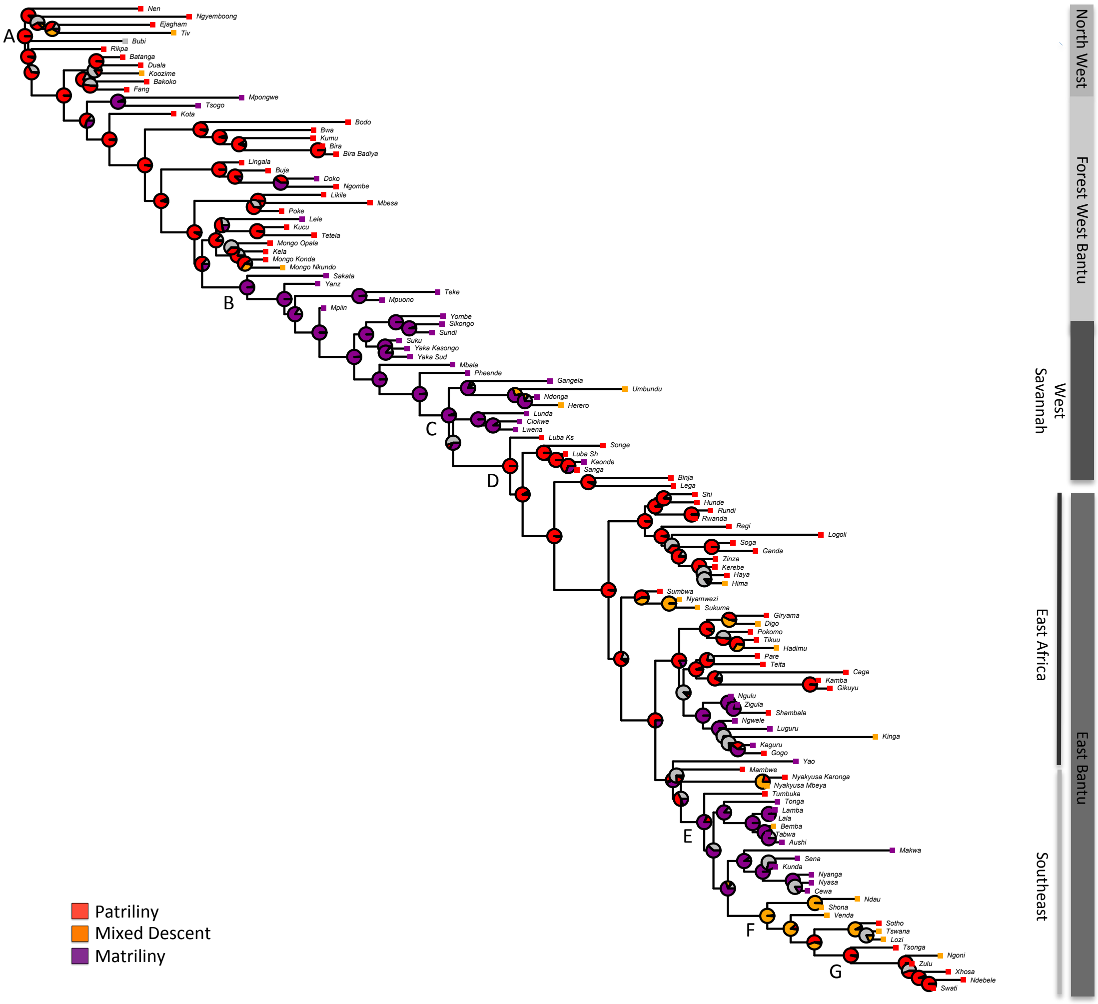

Figure S3. Ancestral reconstruction of descent across Bantu cultures. The pies at each node show the mean probability of each state from the MCMC analysis plotted on the consensus tree. Grey in the pies shows uncertainty that the node exists in the tree sample. Letters (A-G) denote switch points in the state of descent along the expansion route on the phylogeny, for comparison with the residence trait (Figure S4). Cultural groups follow Guthrie (32).

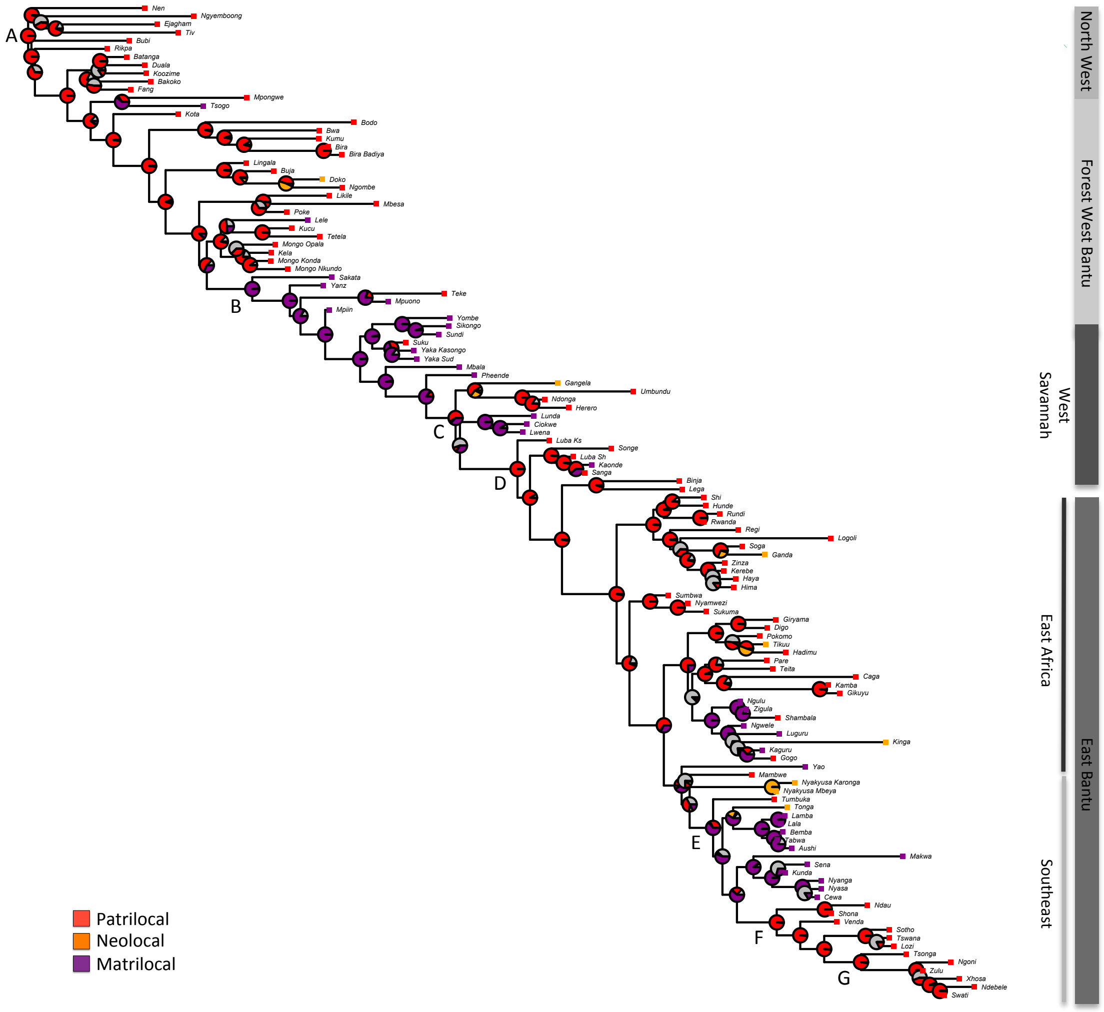

Figure S4 Ancestral reconstruction of residence across Bantu cultures. The pies at each node show the mean probability of each state from the MCMC analysis plotted on the consensus tree. Grey in the pies shows uncertainty that the node exists in the tree sample. Letters (A-G) denote switch points in 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).

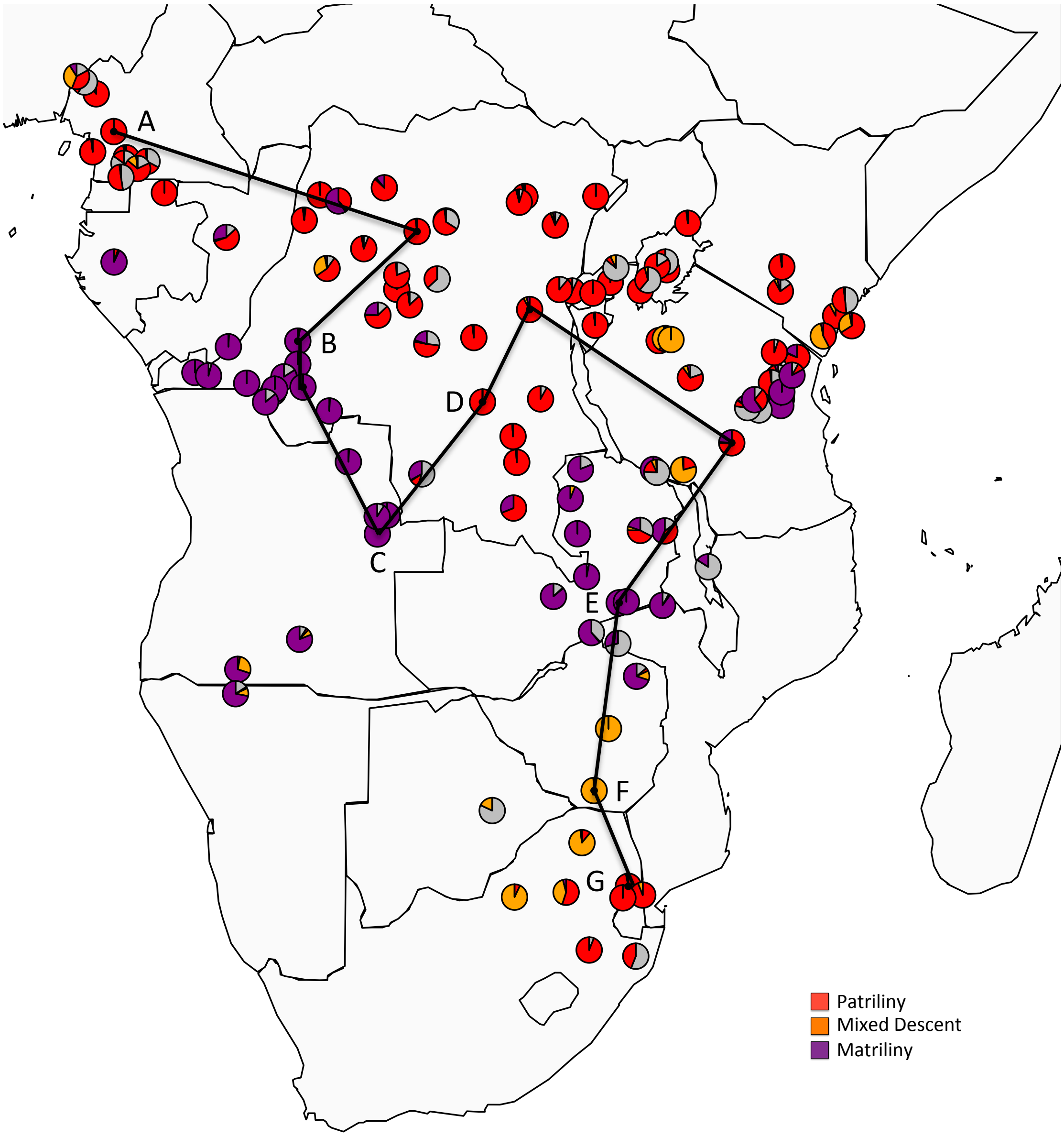

Figure S5 Ancestral reconstruction of descent across Bantu cultures with geographical location and expansion-wave route shown. (Grey in pie shows uncertainty that the node exists from the tree sample). Letters (A-G) denote switch points in the state of descent along the expansion route on the phylogeny, for comparison with the phylogenetic position of switches (Figure 1 & S3) and the switches in the residence trait (Figures S4 & S6).

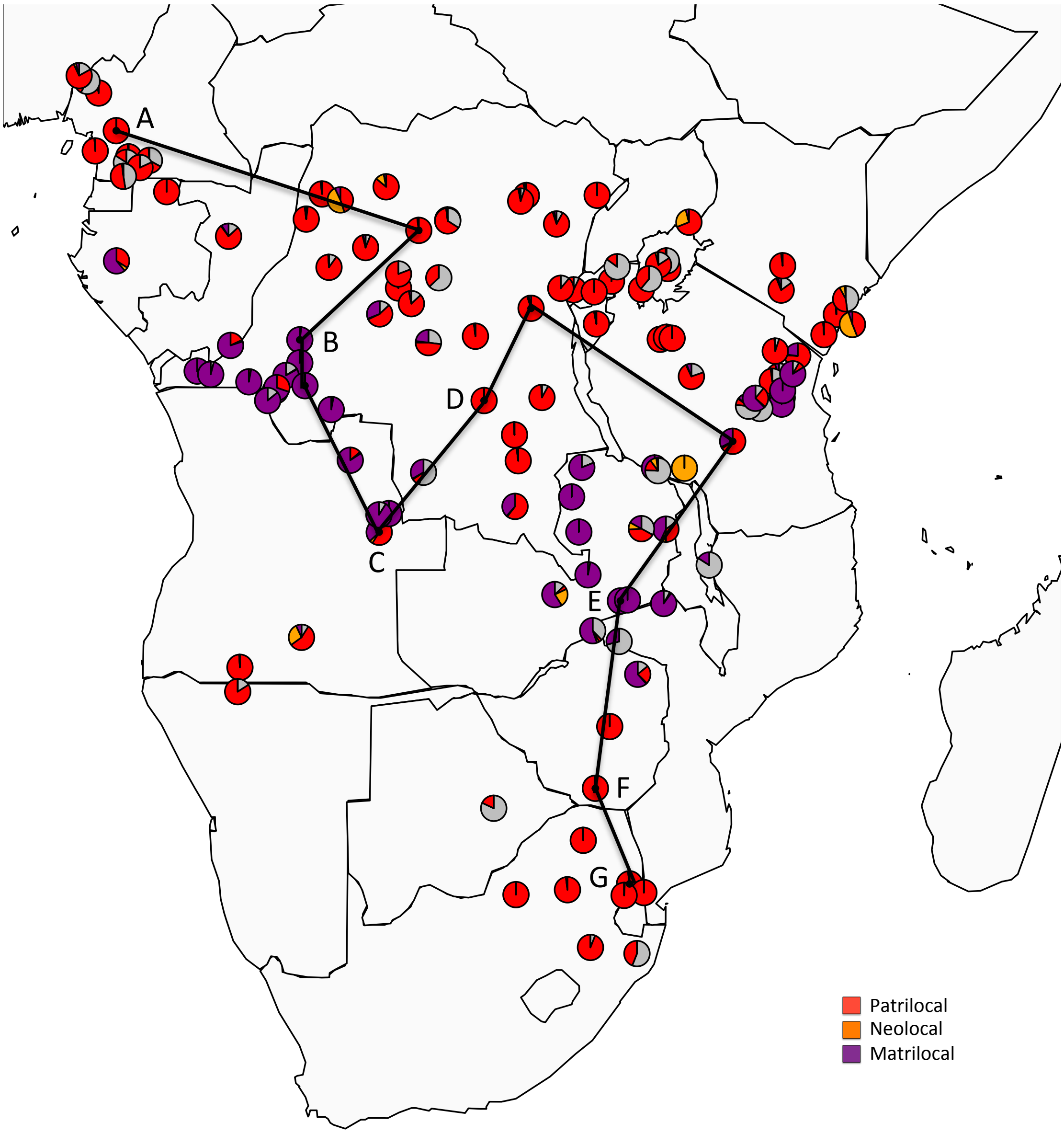

Figure S6 Ancestral reconstruction of residence across Bantu cultures with geographical location and expansion-wave route shown. (Grey in pie shows uncertainty that the node exists from the tree sample). Letters (A-G) denote switch points in the state of residence along the expansion route on the phylogeny, 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 patriliny to matriliny 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.

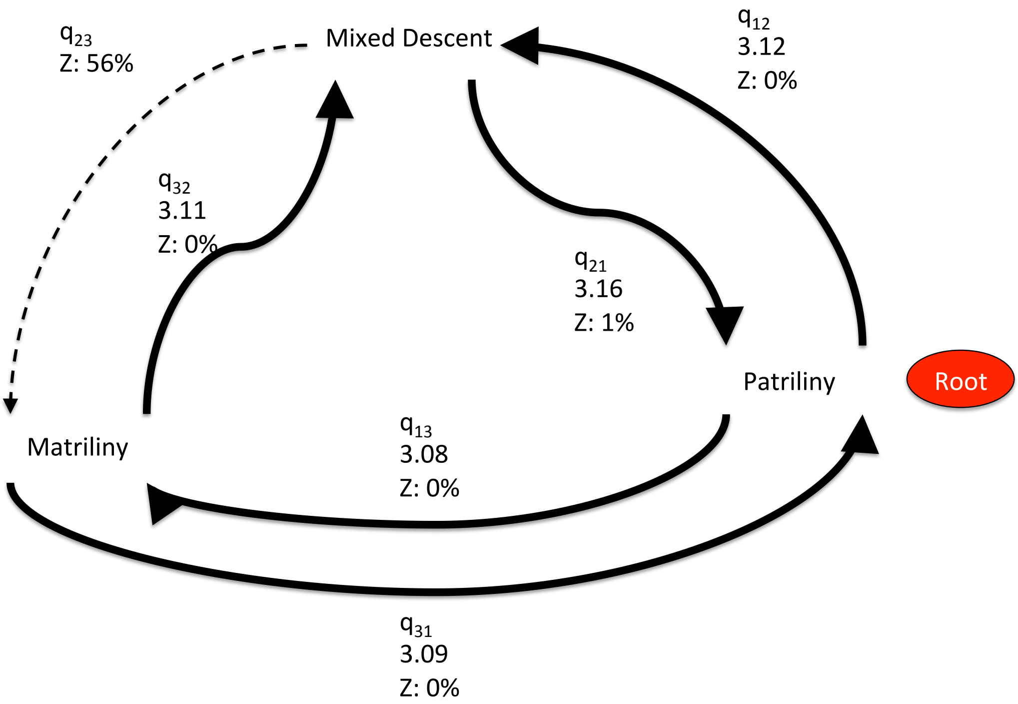

Figure S7 Estimated rate parameters of evolution of descent among Bantu populations under the RJMCMC analysis. Z denotes visits assigned to zero as a proportion of posterior probability distribution. Thickness of arrows reflects proportion of time transition rate is not assigned to zero. Dashed line 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.

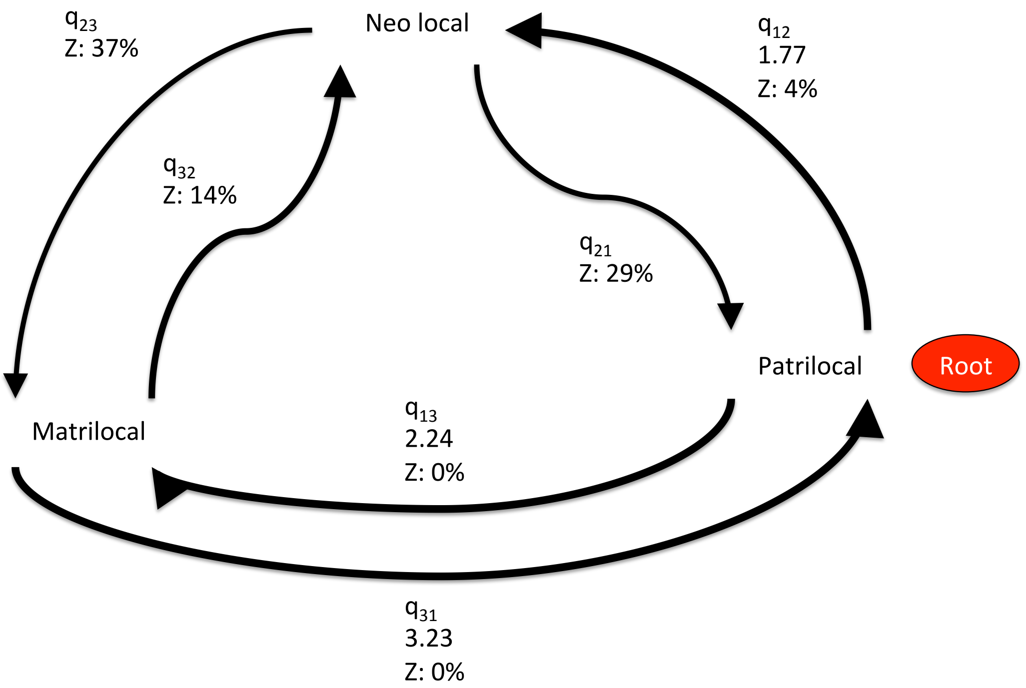

Figure S8 Model of evolution for residence among Bantu populations. Z denotes visits assigned to zero as a proportion of posterior probability distribution. Thickness of arrows reflects proportion of time transition rate is not assigned to zero. Number below rate name (q_{ij}) is the mean transition 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 S4 Comparison of dependent and independent models to test for correlated evolution between traits.

Correlated evolution	**Dependent model**		**Independent model**		**Log_{10} Bayes Factor**
	Log likelihood	S.E.	Log likelihood	S.E.	
Descent - Residence	- 121.45	+/- 0.04	- 128.62	+/- 0.03	3.11

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.

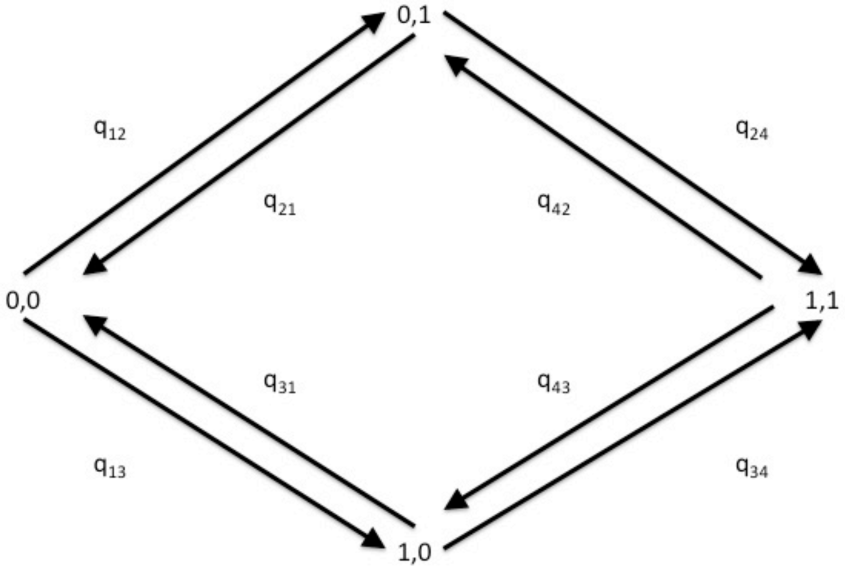

Figure S9 Correlated evolution between two binary traits. There are two routes from the ancestral state (0,0) to the derived state (1,1), either via (0,1) a change in state two first, or (1,0) a change in state one first (it is assumed that the states do not change at the same time over a sufficiently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

Table S5 Comparison of transition rates forced to zero – Patrifocal to Matrifocal.

Model	Log Likelihood	S.E.	Log_{10} Bayes Factor with Model C=0
Transition rates via state C are zero	-138.22	0.211	-
Transition rates via state B are zero	-126.69	0.109	5.01

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.

Model	Log Likelihood	S.E.	Log_{10} Bayes Factor with Model C=0
Transition rates via state C are zero	-132.79	0.241	-
Transition rates via state B are zero	-122.60	0.109	4.43

Supplementary References

1. Parkin R (1997) *Kinship: An Introduction to Basic Concepts* (Blackwell, Oxford).
2. Murdock GP (1949) *Social Structure* (MacMillan, New York).
3. Currie TE, Meade A, Guillon M, & Mace R (2013) Cultural phylogeography of the Bantu Languages of sub-Saharan Africa. *Proceedings of the Royal Society B: Biological Sciences* 280(1762).
4. Fritz SA & Purvis A (2010) Selectivity in Mammalian Extinction Risk and Threat Types: a New Measure of Phylogenetic Signal Strength in Binary Traits. *Conservation Biology* 24(4):1042-1051.
5. Orme C*, et al.* (2011) caper: Comparative Analyses of Phylogenetics and Evolution in R. *URL:* http://cran.r-project.org/web/packages/caper/index.html.
6. R Development Core Team (2008) R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria).
7. Pagel MD, Meade A, & Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology* 53(5):673-684.
8. Pagel MD & Meade A (2006) Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *American Naturalist* 167(6):808-825.
9. Guthrie M (1967-1971) *Comparative Bantu: an introduction to the comparative linguistics and prehistory of the Bantu languages* (Gregg International, Farnborough, UK).
10. Kass RE & Raftery AE (1995) Bayes factors. *Journal of the American Statistical Association* 90(430):773-795.
11. Holden C & Mace R (2003) Spread of cattle led to the loss of matrilineal descent in Africa: a coevolutionary analysis. *Proceedings of the Royal Society of London Series B-Biological Sciences* 270(1532):2425-2433.