

Genetics and Demographic History of the Bantu

Jorge Rocha, *CIBIO/InBIO and Department of Biology, Faculty of Sciences, University of Porto, Porto, Portugal*

Anne-Maria Fehn, *Department of Linguistic and Cultural Evolution, Max-Planck Institute for the Science of Human History, Jena, Germany and CIBIO/InBIO, University of Porto, Porto, Portugal*

Advanced article

Article Contents

- Introduction
- The Migration of Bantu-speaking Peoples out of West Africa
- Migratory Routes
- Admixture with Other Populations

Online posting date: 15th November 2016

Peoples speaking languages of the Bantu family are widespread in sub-Saharan Africa, from the equatorial rainforest to the Cape of Good Hope. Their present-day distribution is the result of a remarkable expansion, which started about 4000–5000 years ago in the borderland between Cameroon and Nigeria. The genetic distances among Bantu-speaking populations are significantly lower than those between the Bantu and other ethnolinguistic groups from Africa, suggesting an actual movement of people, rather than cultural diffusion. However, this genetic homogeneity places a challenge to the reconstruction of their actual routes of dispersal, with existing hypotheses being mainly based on data from linguistics and archaeology. While the indigenous populations they encountered did not provide a major contribution to the Bantu gene pool at large, admixture studies reveal different patterns of social interaction, which range from Bantu dominance in the rainforest to a more levelled exchange in southern Africa.

Introduction

The Bantu family, whose name derives from the shared root *-ntu 'person', is a group of approximately 500 closely related languages that belong to a recently differentiated subbranch of the Niger–Congo phylum of African languages (Figure 1). Presently, the Bantu languages are spoken by about 250 million people across a large geographical area in sub-Saharan Africa (Figure 1). The wide distribution and low degree of internal differentiation

of these languages suggests that their present distribution must be the result of a relatively recent, rapid spread, which came to be known as the Bantu expansion. However, there are many aspects of the history of Bantu languages and Bantu-speaking populations that remain to be clarified and are the object of intense multidisciplinary research, involving contributions from linguistics, ethnography, archaeology and genetics.

Until recently, most studies on the genetics of Bantu-speaking peoples were based on the study of mitochondrial deoxyribonucleic acid (mtDNA) and the nonrecombining region of the Y chromosome (NRY). As both mtDNA and NRY lack recombination, it is relatively easy to infer their phylogenies by reconstructing the order in which mutations accumulate in different lineages. See also: [Mitochondrial Genome: Evolution](#); [Y Chromosome](#); [Chromosome Y](#); [Mitochondrial DNA Polymorphisms](#). The analysis of the geographic distribution of phylogenetically related mtDNA and NRY lineages, known as phylogeography, has provided valuable insights into the genetic relationships between different Bantu-speaking peoples and between Bantu and non-Bantu populations (Pereira *et al.*, 2001; Salas *et al.*, 2002; Berniell-Lee *et al.*, 2009; de Filippo *et al.*, 2011). See also: [Phylogeography](#); [Genetic Diversity in Africa](#). Moreover, owing to their uniparental patterns of inheritance, mtDNA and NRY have been used to explore female and male-specific aspects of Bantu genetic variation. However, these markers represent only a very small fraction of the entire genome and can only offer limited information about the history of populations.

Recent progresses in DNA sequencing and single-nucleotide polymorphism (SNP) array technologies have made it possible to generate genome-wide data based on thousands of autosomal genetic polymorphisms. See also: [Single Nucleotide Polymorphism \(SNP\)](#); [Single Nucleotide Polymorphisms \(SNPs\): Identification and Scoring](#); [HapMap Project](#); [Whole Genome Resequencing and 1000 Genomes Project](#); [High-Throughput Single Nucleotide Polymorphisms Genotyping Technologies](#); [Next Generation Sequencing Technologies and Their Applications](#). The use of these genetic datasets composed of multiple, independently evolving genetic systems in combination with new analytical tools created a unique opportunity to address outstanding questions of population history with an unprecedented resolution (Pugach and Stoneking, 2015). See also:

eLS subject area: Evolution & Diversity of Life

How to cite:

Rocha, Jorge and Fehn, Anne-Maria (November 2016) Genetics and Demographic History of the Bantu. In: eLS. John Wiley & Sons, Ltd: Chichester.

DOI: 10.1002/9780470015902.a0022892

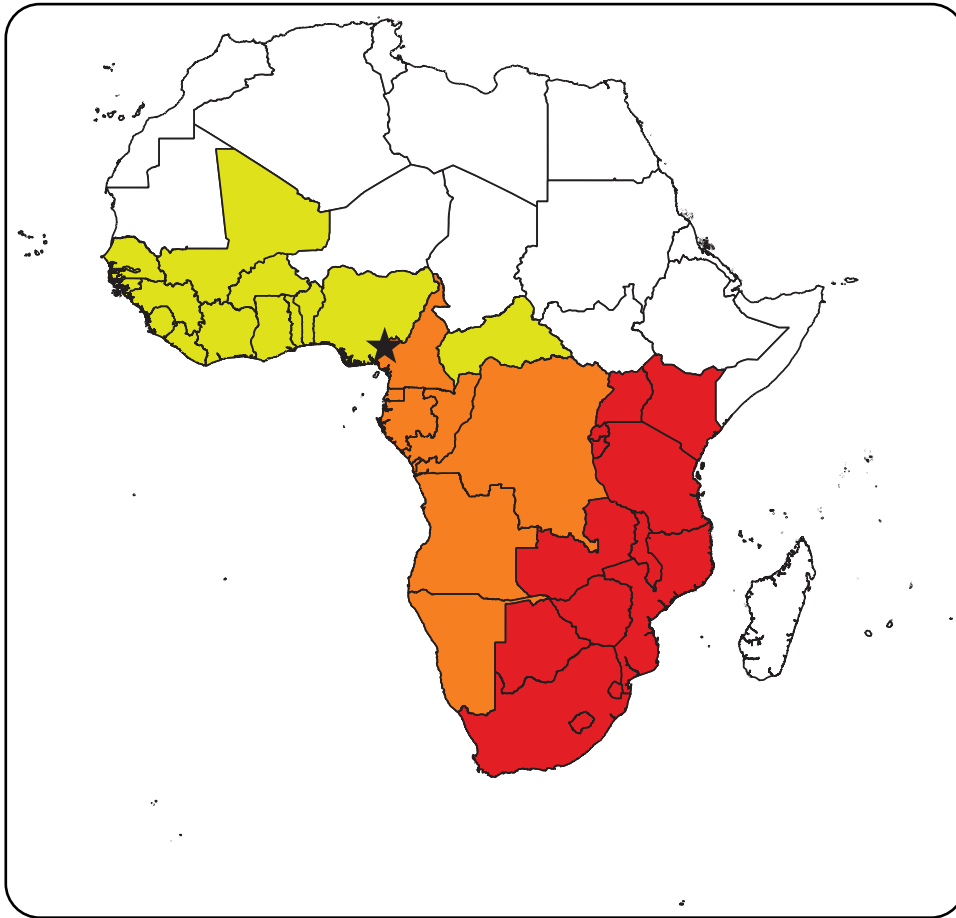


Figure 1 Distribution of non-Bantu Niger–Congo (yellow), West Bantu (Orange) and East Bantu (red) languages. Language groups were assigned according to the affiliation of the most widely distributed languages in the respective countries. The star indicates the approximate origin of the Bantu expansion.

Reconstructing Human History Using Autosomal, Y-Chromosomal and Mitochondrial Markers; Human Relationships Inferred from Genetic Variation

Although the sampling coverage and marker density in Bantu genetic studies is still far from being fully satisfactory, there are many important aspects of the demographic history of Bantu-speaking peoples that have been uncovered with the available mtDNA, NRY and autosomal genome-wide data. In this article, we present a brief outline of the current state of the research on the genetics of Bantu peoples by highlighting these contributions and relating them with the insights provided by other disciplines.

The Migration of Bantu-speaking Peoples out of West Africa

The most obvious evidence for the spread of Bantu languages is, of course, provided by linguistics. Early travellers already noticed that many of the languages spoken in sub-Saharan Africa are similar to each other. Today, we know that this similarity is

due to the recent diversification of all Bantu languages from a common ancestor. The borderland between south-eastern Nigeria and western Cameroon was suggested by Greenberg (1972) to be the original location of the ancestral Bantu language and this idea is now widely accepted. The reasoning behind this assumption is that it is there that Bantu languages are more diverse and meet with their closest relatives of the Niger–Congo family – the so-called Bantoid or Wide Bantu languages (**Figure 1**). However, the linguistic data alone does not tell us whether the spread of Bantu languages and culture actually involved a movement of people or whether it is merely the result of cultural diffusion. **See also: Human Genetics and Languages.** Yet, these two alternatives imply very different predictions about the genetic composition of Bantu-speaking populations from different parts of Africa.

If the spread of Bantu languages was driven only by cultural diffusion and language shift, the different Bantu-speaking populations would not share a distinctive genetic composition and Bantu groups would not be expected to be more similar to each other than to their non-Bantu closest neighbours. These expectations are clearly not met by the available genetic data. For example, Tishkoff *et al.* (2009), using a set of 1300

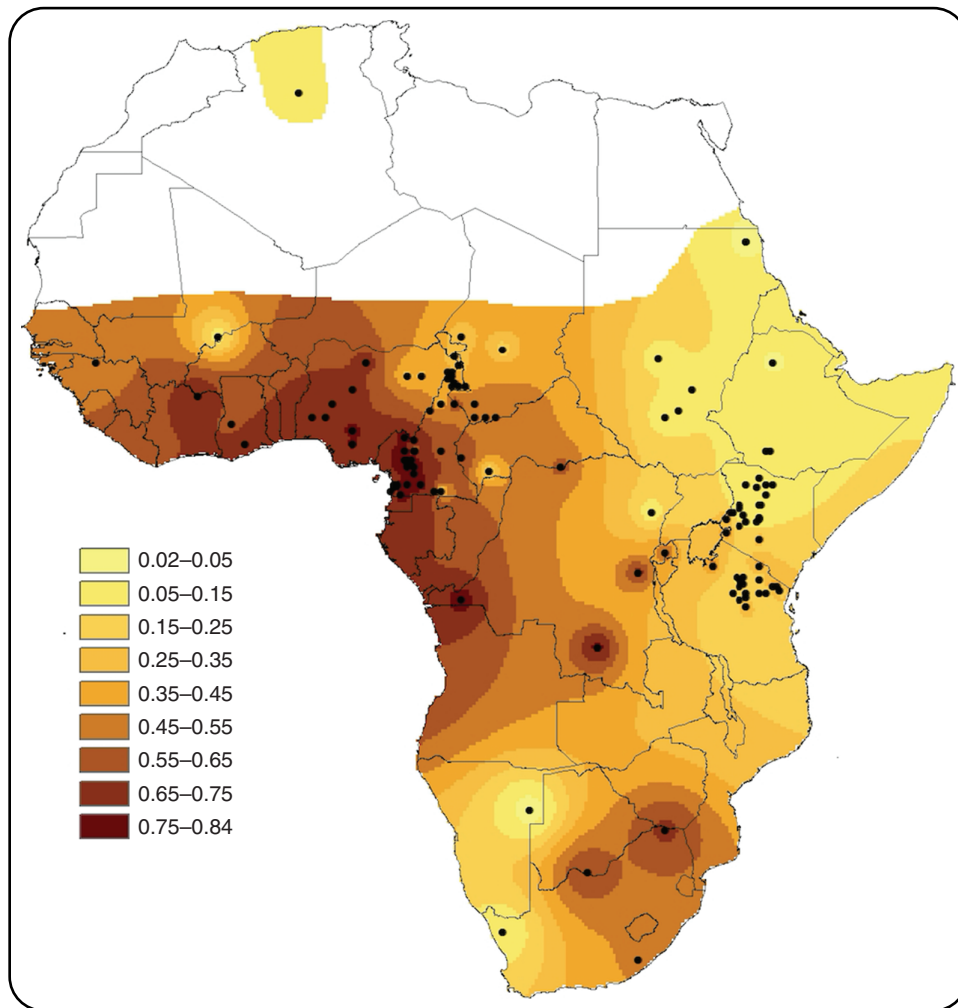


Figure 2 Geographical distribution of the genetic component associated with Niger–Congo-speaking groups. The darker the tone is, the higher is the frequency of the genetic component (see scale). Lighter coloured regions include a comparatively high number of non-Niger–Congo-speaking populations. The map was drawn using data from Tishkoff *et al.* (2009).

autosomal polymorphisms, showed that most Bantu and other Niger–Congo-speaking populations from widely separated geographical areas do share a genetic component that clearly sets them apart from African populations that speak other languages (**Figure 2**). Conversely, Bantu-speaking groups that do not share this genetic component are exceptional. The most prominent case is represented by the genetically and culturally distinct Pygmy hunter-gatherers from the western Congo Basin, who speak Bantu languages probably due to language shift (Bahuchet, 2012). However, these populations are geographically close to and genetically admixed with non-Pygmy Bantu-speaking peoples, indicating that language shift was promoted by the actual migration of these peoples into Pygmy territories (see below).

The genetic unity of Bantu peoples and their close relationship with populations speaking non-Bantu languages within Niger–Congo are additionally reflected in the distribution of uniparental lineages. More than 60% of the NRY lineages of Niger–Congo-speaking peoples belong to a single haplogroup:

E1b1a (Underhill *et al.*, 2001; Rosa *et al.*, 2007; Berniell-Lee *et al.*, 2009; de Filippo *et al.*, 2011). The most frequent maternal lineages found in Bantu populations belong to a more diverse set of haplogroups (mainly L0a, L1c, L2a, L3b and L3e) than observed for the NRY. However, these populations are also closely related on the basis of their mtDNA sequences and trace the bulk of their maternal heritage to western Central Africa (Salas *et al.*, 2002; Beleza *et al.*, 2005; Coelho *et al.*, 2009).

The congruence between the genetic patterns observed with different types of markers has been further confirmed by de Filippo *et al.* (2012), who found that the genetic distances among Bantu-speaking populations based on mtDNA, NRY and autosomal data were all significantly lower than those between the Bantu and other ethnolinguistic groups from Africa.

It is still unclear what the initial triggers for the migration of Bantu peoples out of West Africa were. The beginning of the Bantu expansion is commonly dated to 5000–4000 years ago (Vansina, 1984; Blench, 2006). These dates correlate with

the appearance in western Central Africa of archaeological innovations associated with new subsistence strategies that were probably made possible by climate-induced openings in the rain forest (Bostoen *et al.*, 2015). For a long time, it was believed that the Bantu expansion was directly linked to the advent and spread of agriculture (Diamond and Bellwood, 2003). However, the archaeological record does not seem to support this assumption, and it is likely that the early migrations of Bantu speakers were rather associated with a more efficient use of natural resources (Blench, 2006; Bostoen *et al.*, 2015). In any case, the later addition of agriculture and iron working to these subsistence strategies must have had an important impact on the spread of the Bantu by affording them a substantial technological advantage over local foragers (Diamond and Bellwood, 2003). This impact is illustrated by the early arrival of Bantu peoples possessing knowledge on both crop farming and iron working to KwaZulu-Natal, in the southern tip of Africa, only about 2000 years after the first population movements in West Africa (Bostoen *et al.*, 2015).

Several studies have attempted to provide a genetic-based dating for the Bantu expansion using two main types of approaches. The first approach estimates the time needed to generate the diversity presently observed in specific mtDNA or NRY lineages that are considered to be associated with Bantu-speaking populations (Pereira *et al.*, 2001; Salas *et al.*, 2002; Zhivotovsky *et al.*, 2004; Berniell-Lee *et al.*, 2009; de Filippo *et al.*, 2011; Soares *et al.*, 2012). The second approach is based on explicit models of population behaviour and uses mtDNA, NRY or autosomal data to infer the divergence times between Bantu groups scattered across different geographical regions (Coelho *et al.*, 2009; Alves *et al.*, 2011; Montano *et al.*, 2011; Li *et al.*, 2014). Both types of methods produced time estimates that are consistent with archaeology-based estimates. Moreover, some of these works have explicitly investigated population size changes and found signals of population growth that slightly precede or overlap with the beginning of the Bantu dispersals (Pilkington *et al.*, 2008; Coelho *et al.*, 2009; Alves *et al.*, 2011; Gignoux *et al.*, 2011; Batini *et al.*, 2011; Montano *et al.*, 2011; Li *et al.*, 2014).

Interestingly, the comparison of demographic data obtained in Bantu populations for NRY and mtDNA suggests that males had lower effective population sizes than females (Destro-Bisol *et al.*, 2004; Wood *et al.*, 2005; Pilkington *et al.*, 2008; Coelho *et al.*, 2009; Verdu *et al.*, 2013). **See also: Effective Population Size.** This finding, which is reflected in the differences between mtDNA and NRY lineage diversity (see above), can be explained by the generalised practice of polygyny among farming Bantu populations (Destro-Bisol *et al.*, 2004). By increasing the differences in male reproductive success, this practice accelerates the rate of lineage extinction and leads to a strong reduction of the NRY genetic variation. In contrast, females maintain a more even distribution of fertility, preserving the mtDNA diversity of the ancestral population.

Patrilocality, the displacement of married women to the place of residence of their husbands' family, is another sociocultural rule that might have shaped sex-specific patterns of genetic diversity among Bantu populations by promoting larger mobility in women than in men. These differences are reflected in the levels of genetic differentiation among Bantu groups, which are more

elevated for NRY than for mtDNA (Destro-Bisol *et al.*, 2004; Coelho *et al.*, 2009; Verdu *et al.*, 2013).

Migratory Routes

Most views about the major dispersal routes followed by Bantu speakers can be divided into two main models that were largely influenced by the interpretation of linguistic data (Pakendorf *et al.*, 2011) (**Figure 3**). According to one of the models, known as the early-split hypothesis, the Bantu languages can be primarily divided into east and west monophyletic branches (Currie *et al.*, 2013) (**Figure 3a**). In this model, the eastern branch is usually linked to a movement of Bantu-speaking peoples along the northern fringes of the rainforest into the Great Lakes and then southwards to southeast Africa (Phillipson, 1977). The western branch is associated with an initial southward move, followed by dispersals across the western part of subequatorial Africa. According to the alternative model, known as the late-split hypothesis, eastern Bantu diverged from its western linguistic relatives only after Bantu-speaking peoples had crossed the rainforest (**Figure 3b**).

Recently, Currie *et al.* (2013) built language phylogenies that support the late-split model by showing that the eastern subgroup is a relatively late offshoot of Bantu languages spoken in the western half of Africa. Moreover, by inferring the geographical location of ancestral Bantu societies, they reconstructed the most likely dispersal scenarios that explain the uncovered phylogenetic relationships among different extant languages.

The views about the migratory paths associated with the linguistic hypotheses are also motivated by different interpretations about the capacity of the Bantu to cross the rainforest (**Figure 3**). To explore the effects of different environments on the spread of Bantu peoples, Russell *et al.* (2014) analysed the available archaeological evidence on the arrival dates of early farming into Bantu-speaking regions of Africa. Their results indicate that forest habitats can indeed be an obstacle to the penetration of agriculturalist populations and emphasise the role of rivers, coastlines and a savannah corridor through the rainforest in facilitating the southward migration of Bantu groups. However, the compiled archaeological data was still insufficient to indicate if the dispersal routes involved an early or late split between eastern and western Bantu-speaking populations. Grollemund *et al.* (2015), using a dated language phylogeny to analyse the pace of the Bantu migrations, have also emphasised the importance of savannah corridors emerging around 4000 years ago in the penetration of Bantu peoples into the rainforest (**Figure 3b**). Moreover, they found strong evidence for the late-split model and inferred dispersal routes that are close to those favoured by Currie *et al.* (2013), including a major passage through the savanna corridor without an exploration of coastal routes.

The available genetic evidence has not yet provided a robust discrimination of the alternative migratory models. Recently, de Filippo *et al.* (2012) formally tested the early and late-split hypotheses by correlating genetic and linguistic data with the geographical distances between populations implied by the different dispersal models. They found that migration distances associated with the early-split model were the worst predictors of the current levels of linguistic and genetic differentiation among

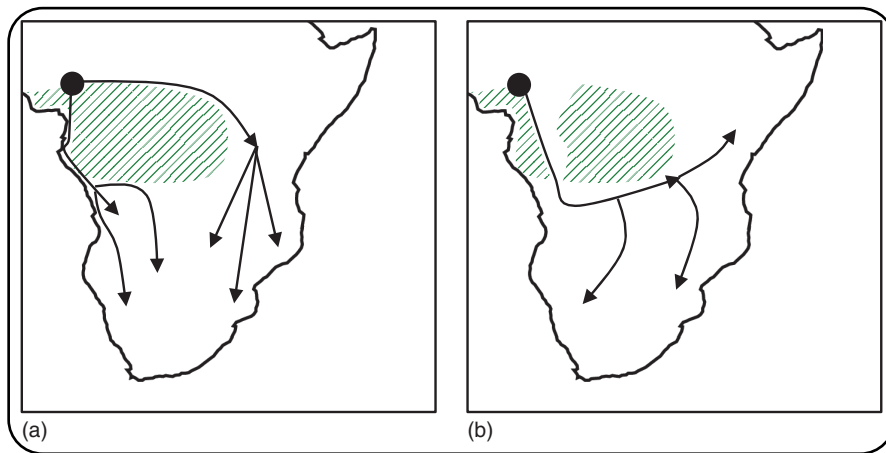


Figure 3 Models of Bantu dispersal. (a) Early split; (b) Late split. Based on information from maps published in Pakendorf *et al.* (2011); Russell *et al.* (2014), and Grollemund *et al.* (2015).

the tested Bantu populations. However, while the late-split model showed a relatively good fit to the data, it was the shortest distance between populations that displayed the highest correlation with the linguistic and genetic evidence. These results were interpreted as an indication that genetic admixture and linguistic borrowing between neighbouring Bantu communities could have weakened the signal of their historical migrations.

Other genetic works focused on more restricted aspects of the migratory pathways of Bantu populations using different kinds of molecular markers. For example, Alves *et al.* (2011), using a set of 14 autosomal polymorphisms, observed a high genetic homogeneity between Bantu groups from Angola and Mozambique. This homogeneity was found hard to explain according to the early-split model, which places the studied populations at two extreme edges of the Bantu migrations. However, owing to the limited number of markers used in this study, it is not clear whether the observed homogeneity is simply due to the low resolution of the genetic data. More recently, Li *et al.* (2014) reanalysed part of the genetic data from Tishkoff *et al.* (2009), consisting of 717 autosomal markers, to explicitly investigate alternative hypotheses about routes of dispersal and genetic relationships of eastern Bantu populations. They found that eastern Bantu-speaking groups from Kenya, Tanzania and South Africa share a more recent common ancestor with each other than with other Bantu groups. However, this study could not discriminate between the early and the late-split models, as both hypotheses recognise that eastern Bantu populations are monophyletic (**Figure 3**). Moreover, in spite of the large number of loci analysed, the grouping of eastern Bantu populations received only marginal support compared to other models of population bifurcation.

A common genetic component to Bantu groups from eastern and south-eastern Africa has recently been revealed using a dataset of 1747 SNPs (González-Santos *et al.*, 2015). This study confirms the close genetic relationship between eastern Bantu speakers and highlights the possibility of disclosing substructure within Bantu-speaking groups that cannot be ascribed to

admixture with non-Bantu populations. Future work increasing both the range of sampled populations and the resolution of the genetic analyses will be needed to further explore the genetic relationships among different Bantu groups and infer the dispersal models that best explain them.

Admixture with Other Populations

In principle, the mere observation that geographically distant Bantu communities share a discernable genetic component implies that the incorporation of genetic contributions from local peoples cannot be a pervasive feature of the Bantu expansions. Otherwise, the original genetic characteristics of the Bantu peoples would have been diluted and the aforementioned evidence for a population movement would have been lost (Diamond and Bellwood, 2003).

However, the study of the interactions between migrating Bantu peoples and indigenous groups is very important for assessing the impact of admixture and sociocultural factors on the local differentiation of Bantu groups, as well as on the genetic composition of resident populations.

Presently, the best documented cases of admixture between Bantu and non-Bantu populations are provided by studies that were undertaken in the Central African rainforest and the southern African Kalahari basin.

Several studies on the peoples of the western Central African rainforest have consistently shown that Bantu groups had a substantial impact on the genetic composition of Pygmy hunter-gatherers (Berniell-Lee *et al.*, 2009; Verdu *et al.*, 2009, 2013; Patin *et al.*, 2014). For example, Patin *et al.* (2014) using genome-wide data from over 300 000 SNPs estimated that some Pygmy forager groups display up to 50% of Bantu genomic ancestry. Moreover, the analysis of uniparental markers has shown that Bantu-to-Pygmy admixture was mostly due to matings between Bantu males and Pygmy females who remained at or returned to their own villages with their children (Berniell-Lee

et al., 2009; Verdu *et al.*, 2013; Patin *et al.*, 2014). Conversely, introgression of Pygmy genetic material into Bantu-speaking groups was found to be low and limited to mtDNA lineages carried by Pygmy females (Destro-Bisol *et al.*, 2004; Quintana-Murci *et al.*, 2008; Batini *et al.*, 2011; Verdu *et al.*, 2013; Patin *et al.*, 2014). These patterns of male-mediated asymmetric gene flow from Bantu to Pygmy groups have been interpreted in the context of nonleveled social interactions in which foraging groups are discriminated by their dominant agriculturalist Bantu neighbours (Verdu *et al.*, 2013). A striking cultural consequence of this background was the shift of hunter-gatherer Pygmy communities to the languages of incoming Bantu groups (Bahuchet, 2012; Verdu and Destro-Bisol, 2012).

Southern Africa provides a quite different example. The peoples who inhabited the region before the arrival of Bantu-speaking agriculturalists, generally designated 'Khoisan', are often

considered to be a homogenous group on the basis of their highly distinctive click languages. However, these groups are in fact very diverse and the click languages from southern Africa are now thought to belong to three independent lineages: Tuu, Kx'a and Khoe-Kwadi (Güldemann, 2014). The populations speaking languages from the Khoe-Kwadi family, in particular, are notorious for their biological and cultural variation. Unlike Tuu- and Kx'a-speaking peoples, who are preferentially hunter-gatherers, Khoe-Kwadi groups include both pastoralist and foraging communities and often present important levels of admixture with Bantu. For example, the inferences from Pickrell *et al.* (2014), based on a genome-wide array of about 550 000 autosomal SNPs, show that the proportion of Bantu ancestry in most Khoe-Kwadi speakers is consistently higher than in other indigenous peoples of southern Africa (**Figures 4 and 5a**). In the case of the Damara of Namibia, the Bantu proportion

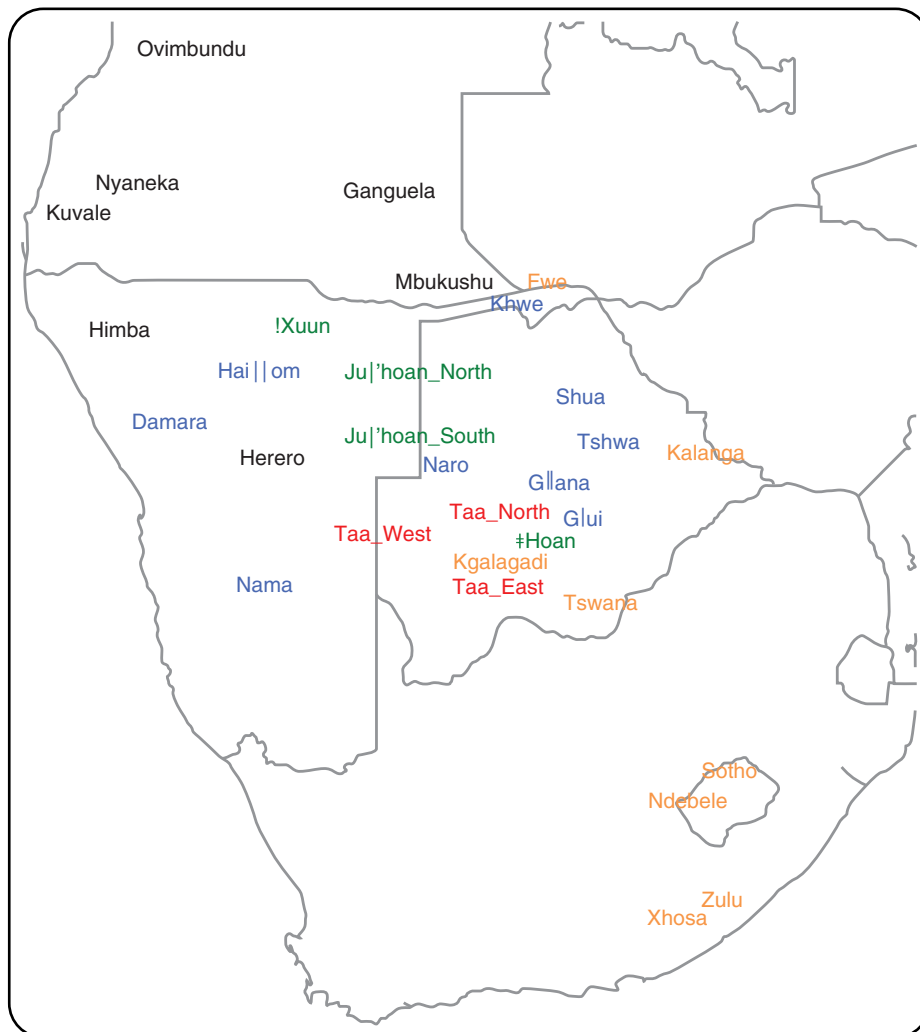


Figure 4 Locations of sampled populations referred to in **Figure 5**. Colours indicate different language groups. Black: West Bantu; Orange: East Bantu; Green: Kx'a; Red: Tuu; Blue: Khoe-Kwadi. Geographical location of populations based on Barbieri *et al.* (2014a,b); Pickrell *et al.* (2014), and Marks *et al.* (2015).

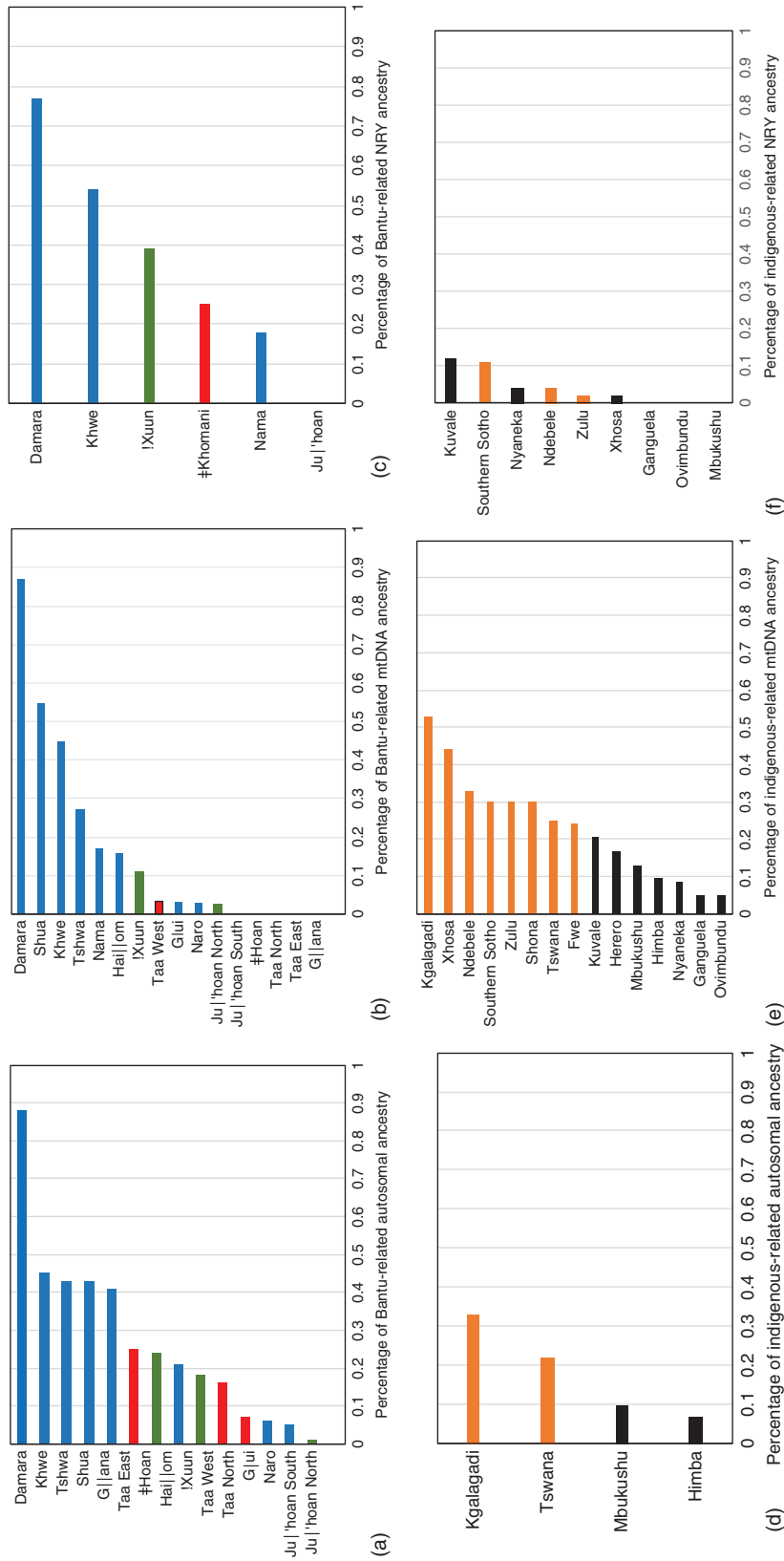


Figure 5 Admixture between Bantu-speaking and indigenous populations of southern Africa. (a–c) Inferred proportions of Bantu-related ancestry among indigenous populations. (d–f) Inferred proportions of indigenous-related ancestry among Bantu populations. The geographical positions of sampled groups are shown in **Figure 4**. Graphic bars are coloured according to language groups as in **Figure 4**. Graphics were generated using data from Barbieri *et al.* (2014a,b); Pickrell *et al.* (2014); Marks *et al.* (2015).

is so elevated (80%) that it is likely this group has a Bantu origin, but shifted its language to Khoe-Kwadi. Remarkably, the results from the autosomal markers are highly correlated with the available mtDNA data, indicating that Bantu females actually migrated to Khoe-Kwadi-speaking populations (**Figure 5b**). This migration may be interpreted as a sign of more levelled interactions, as Bantu females would not be expected to move into Khoe-Kwadi communities if the Bantu were dominant. The available data on NRY is not as detailed as for autosomes or mtDNA, but suggests that Kx'a- and Tuu-speaking groups, unlike the Khoe-Kwadi peoples, might have acquired their Bantu genetic component mostly through males, like in the case of Bantu-Pygmy interactions (**Figure 5c**). The reasons for a more intense and levelled interaction between the Bantu- and the Khoe-Kwadi-speaking populations are still poorly understood. According to Güldemann (2008), the Khoe-Kwadi languages were introduced into southern Africa by intruding pre-Bantu pastoralist peoples that originated in East Africa and encroached into regions previously inhabited by Kx'a- and Tuu-speaking hunter-gatherers. Thus, it is plausible that a pastoralist lifestyle of the Khoe-Kwadi promoted more egalitarian relations with incoming Bantu farmers. However, it should be noted that most Khoe-Kwadi groups are presently foragers. The only Khoe-Kwadi population that exhibits a pastoralist subsistence is the Nama from Namibia, who do not present exceptional levels of Bantu admixture (**Figures 4 and 5b, c**).

Admixture in southern Africa is of course also reflected in the patterns of introgression of indigenous genetic material into the incoming Bantu populations (**Figures 5d**), although in this case it is difficult to distinguish the respective contributions of Kx'a, Tuu or Khoe-Kwadi groups (Barbieri *et al.*, 2014b). This introgression is obviously female mediated as the frequency of indigenous-related mtDNA lineages in southern Bantu groups is much higher than NRY (**Figure 5e and f**). However, unlike in the Bantu-Pygmy admixture setting, it seems clear that the indigenous females mating with the Bantu did not return to their original communities as their mtDNA lineages were kept in the Bantu-speaking groups. Interestingly, some of these mtDNA lineages can only be found in Bantu populations, who thus came to preserve part of the genetic legacy of vanished indigenous groups (Barbieri *et al.*, 2013a). Moreover, some interactions have even promoted the acquisition of click phonemes in a number of Bantu languages, instead of leading to language replacement among the resident peoples (Barbieri *et al.*, 2013b). Groups such as the Zulu, Ndebele and Xhosa, who speak Bantu languages with clicks, are actually among the Bantu populations with the highest frequencies of indigenous mtDNA lineages (**Figures 4 and 5e**).

References

Alves I, Coelho M, Damasceno A, *et al.* (2011) Genetic homogeneity across Bantu-speaking groups from Mozambique and Angola challenges early split scenarios between East and West Bantu populations. *Human Biology* **83** (1): 13–38.

Bahuchet S (2012) Changing language, remaining Pygmy. *Human Biology* **84** (1): 11–43.

Barbieri C, Vicente M, Rocha J, *et al.* (2013a) Ancient substructure in early mtDNA lineages of southern Africa. *American Journal of Human Genetics* **92** (2): 285–292.

Barbieri C, Butthof A, Bostoen K and Pakendorf B (2013b) Genetic perspectives on the origin of clicks in Bantu languages from southwestern Zambia. *European Journal of Human Genetics* **21**: 430–436.

Barbieri C, Güldemann T, Naumann C, *et al.* (2014a) Unraveling the complex maternal history of Southern African Khoisan populations. *American Journal of Physical Anthropology* **153**: 435–448.

Barbieri C, Vicente M, Oliveira S, *et al.* (2014b) Migration and interaction in a contact zone: mtDNA variation among Bantu-speakers in southern Africa. *PLoS One* **9** (6): e99117.

Batini C, Lopes J, Behar DM, *et al.* (2011) Insights into the demographic history of African Pygmies from complete mitochondrial genomes. *Molecular Biology and Evolution* **28** (2): 1099–1110.

Beleza S, Gusmão L, Amorim A, Carracedo A and Salas A (2005) The genetic legacy of western Bantu migrations. *Human Genetics* **117**: 366–375.

Berniell-Lee G, Calafell F, Bosch E, *et al.* (2009) Genetic and demographic implications of the Bantu expansion: insights from human paternal lineages. *Molecular Biology and Evolution* **26** (7): 1581–1589.

Blench R (2006) *Archaeology, Language, and the African Past*. Lanham, MD: Altamira Press.

Bostoen K, Clist B, Doumenge C, *et al.* (2015) Middle to Late Holocene paleoclimatic change and the early Bantu expansion in the rain forests of West Central-Africa. *Current Anthropology* **56** (3): 354–384.

Coelho M, Sequeira F, Luiselli D, Beleza S and Rocha J (2009) On the edge of the Bantu expansions: mtDNA, Y chromosome and lactase persistence genetic variation in southwestern Angola. *BMC Evolutionary Biology* **9**: 80.

Currie TE, Meade A, Guillon M and Mace R (2013) Cultural phylogeography of the Bantu languages of sub-Saharan Africa. *Proceedings of the Royal Society B* **280**: 1762.

Destro-Bisol G, Donati F, Cia V, Boschi I, *et al.* (2004) Variation of female and male lineages in sub-Saharan populations: the importance of sociocultural factors. *Molecular Biology and Evolution* **21**: 1673–1682.

Diamond J and Bellwood P (2003) Farmers and their languages: the first expansions. *Science* **300**: 597–603.

de Filippo C, Barbieri C, Whitten M, Sununguko W, *et al.* (2011) Y-chromosomal variation in sub-Saharan Africa: insights into the history of Niger-Congo groups. *Molecular Biology and Evolution* **28** (3): 1255–1269.

de Filippo C, Bostoen K, Stoneking M and Pakendorf B (2012) Bringing together linguistic and genetic evidence to test the Bantu expansion. *Proceedings of the Royal Society B* **279**: 1741.

Gignoux CR, Henn BM and Mountain JL (2011) Rapid, global demographic expansions after the origins of agriculture. *Proceedings of the National Academy of Sciences of the United States of America* **108** (15): 6044–6049.

González-Santos M, Montinaro F, Oosthuizen O, *et al.* (2015) Genome-wide analysis of southern African populations provides new insights into the dispersal of Bantu-speaking groups. *Genome Biology and Evolution* **79** (9): 2560–2568.

Greenberg JH (1972) Linguistic evidence regarding Bantu origins. *The Journal of African History* **13** (2): 189–216.

Grollemund R, Branford S, Bostoen K, *et al.* (2015) Bantu expansion shows that habitat alters the route and pace of human dispersals.

- Proceedings of the National Academy of Sciences of the United States of America* **112** (43): 13296–13301.
- Güldemann T (2008) A linguist's view: Khoe-Kwadi speakers as the earliest food-producers of southern Africa (In: Sadr K and Fauvelle-Aymar FX (eds) *Khoekhoe and the Earliest Herders in Southern Africa*). *Southern African Humanities* **20**: 93–132.
- Güldemann T (2014) “Khoisan” linguistic classification today. In: Güldemann T and Fehn AM (eds) *Beyond ‘Khoisan’: Historical Relations in the Kalahari Basin*, pp. 1–41. Amsterdam: John Benjamins.
- Li S, Schlebusch C and Jakobsson M (2014) Genetic variation reveals large-scale population expansion and migration during the expansion of Bantu-speaking peoples. *Proceedings of the Royal Society B* **281**: 1793.
- Marks SJ, Montinaro F, Levy H, *et al.* (2015) Static and moving frontiers: the genetic landscape of southern African Bantu-speaking populations. *Molecular Biology and Evolution* **32** (1): 29–43.
- Montano V, Ferri G, Marcari V, *et al.* (2011) The Bantu expansion revisited: a new analysis of Y chromosome variation in Central Western Africa. *Molecular Ecology* **20**: 2693–2708.
- Patin E, Siddle KJ, Laval G, *et al.* (2014) The impact of agricultural emergence on the genetic history of African rainforest hunter-gatherers and agriculturalists. *Nature Communications* **5**: 3163.
- Pereira L, Macaulay V, Torroni A, *et al.* (2001) Prehistoric and historic traces in the mtDNA of Mozambique: insights into the Bantu expansions and the slave trade. *Annals of Human Genetics* **65**: 439–458.
- Pakendorf B, Bostoen K and de Filippo C (2011) Molecular perspectives on the Bantu expansion: a synthesis. *Language Dynamics and Change* **1** (1): 50–58.
- Phillipson DW (1977) *The Later Prehistory of Eastern and Southern Africa*. London: Heinemann.
- Pickrell JK, Patterson N, Loh PR, *et al.* (2014) Ancient west Eurasian ancestry in southern and eastern Africa. *Proceedings of the National Academy of Sciences of the United States of America* **111** (7): 2632–2637.
- Pilkington MM, Wilder JA, Mendez FL, *et al.* (2008) Contrasting signatures of population growth for mitochondrial DNA and Y chromosomes among human populations in Africa. *Molecular Biology and Evolution* **25** (3): 517–525.
- Pugach I and Stoneking M (2015) Genome-wide insights into the genetic history of human populations. *Investigative Genetics* **6**: 6.
- Quintana-Murci L, Quach H, Harmant C, *et al.* (2008) Maternal traces of deep common ancestry and asymmetric gene flow between Pygmy hunter-gatherers and Bantu-speaking farmers. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 1596–1601.
- Rosa A, Ornelas C, Jobling MA, Brehm A and Villems R (2007) Y-chromosomal diversity in the population of Guinea Bissau: a multiethnic perspective. *BMC Evolutionary Biology* **7**: 124.
- Russell T, Silva F and Steele J (2014) Modelling the spread of farming in the Bantu-speaking regions of Africa: an archaeology-based phylogeography. *PLoS One* **9** (1): e87854.
- Salas A, Richards M, De la Fe T, *et al.* (2002) The making of the African mtDNA landscape. *American Journal of Human Genetics* **71** (5): 1082–1111.
- Soares P, Alshamali F, Pereira JB, *et al.* (2012) The expansion of mtDNA haplogroup L3 within and out of Africa. *Molecular Biology and Evolution* **29** (3): 915–927.
- Tishkoff SA, Reed FA, Friedlaender FR, *et al.* (2009) The genetic structure and history of Africans and African Americans. *Science* **324**: 1035–1044.
- Underhill PA, Passarino G, Lin AA, *et al.* (2001) The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Annals of Human Genetics* **65** (1): 43–62.
- Vansina J (1984) Western Bantu expansion. *The Journal of African History* **25**: 129–145.
- Verdu P, Austerlitz F, Estoup A, *et al.* (2009) Origins and genetic diversity of pygmy hunter-gatherers from Western Central Africa. *Current Biology* **19**: 312–318.
- Verdu P and Destro-Bisol G (2012) African Pygmies, what's behind a name? *Human Biology* **84** (1): 1–10.
- Verdu P, Becker NSA, Froment A, *et al.* (2013) Sociocultural behavior, sex-biased admixture and effective population sizes in Central African Pygmies and non-Pygmies. *Molecular Biology and Evolution* **30** (4): 918–937.
- Wood ET, Stover DA, Ehret C, *et al.* (2005) Contrasting patterns of Y chromosome and mtDNA variation in Africa: evidence for sex-biased demographic processes. *European Journal of Human Genetics* **13**: 867–876.
- Zhivotovsky LA, Underhill PA, Cinnioğlu C, Kayer M, *et al.* (2004) The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. *American Journal of Human Genetics* **74** (1): 50–61.

Further Reading

- Campbell MC and Tishkoff S (2010) The evolution of human genetic and phenotypic variation in Africa. *Current Biology* **20** (4): 166–173.
- Heyer E and Rocha J (2015) Sub-Saharan Africa: human genetics. In: Bellwood P (ed) *The Global Prehistory of Human Migration*, pp. 115–124. Oxford: Wiley Blackwell.
- Newman JL (1995) *The Peopling of Africa: A Geographic Interpretation*. New Haven/London: Yale University Press.
- Pakendorf B (2015) Historical linguistics and molecular anthropology. In: Bownen C and Evans B (eds) *The Routledge Handbook of Historical Linguistics*, pp. 627–641. Oxon/New York: Routledge.
- Pickrell JK, Patterson N, Barbieri C, *et al.* (2012) The genetic prehistory of southern Africa. *Nature Communications* **3**: 1143.
- Schlebusch CM, Skoglund P, Sjödin P, *et al.* (2012) Genomic variation in seven Khoe-San groups reveals adaptation and complex African history. *Science* **338** (6105): 374–379.