Title

The power of contemporary African DNA: Exploring models of human evolution and health in Africa

Journal

South African Journal of Science, 120(1/2)

ISSN

0038-2353

Authors

Möller, Marlo

Hoal, Eileen

Henn, Brenna M

Publication Date

2024

DOI

10.17159/sajs.2024/17145

Peer reviewed

Title:
The Power of Contemporary African DNA: Exploring Models of Human Evolution and Health in Africa

Running head:
Exploring Models of Human Evolution and Health in Africa

Keywords:
Human evolution
Genetic diversity
African genomes
Health implications
Population genomics

UN Sustainable Development Goals:
Good Health and Wellbeing

Author names and ORCIDs:
Marlo Möller[1] : https://orcid.org/0000-0002-0805-6741
Eileen Hoal[1] : https://orcid.org/0000-0002-6444-5688
Brenna M. Henn[2,3] : https://orcid.org/0000-0003-4998-287X

Affiliations:
[1] DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research, South African Medical Research Council Centre for Tuberculosis Research; Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town.
[2] Department of Anthropology, University of California, Davis, CA, USA.
[3] Genome Center, University of California, Davis, CA, USA.

Corresponding author name and email:
Marlo Möller, marlom@sun.ac.za

Significance

It is generally accepted that humans evolved in Africa, but several opposing conceptual models representing our origins have been proposed. Our study sheds light on the divergence of human populations on the African continent and challenges traditional models, suggesting a new framework - represented by a tangled vine with offshoots - where stem populations separated, but continually exchanged genetic material. This work would not have been possible without sequencing the most genetically diverse human genomes in the world - contemporary African DNA is not only key to understanding deep human history, but is central to answering other health-related questions.

Genetic models of human evolution

The study of human evolution has been a subject of great interest as we seek to understand the origins and history of our species. One of the fundamental organising concepts in biology is the “tree of life”, which was the predominant model of human evolution for many years. This model of human population divergence from a single ancestral population in Africa, was supported by the

existing genetic data, but was difficult to align with the evidence of *Homo sapiens* fossils and archaeological sites across the continent [1].

Fossils and archaeological records indicate the presence of anatomically modern humans across Africa between 300,000 and 100,000 years ago. Several key fossils, such as those found at Jebel Irhoud in Morocco, Herto in Ethiopia, and Klasies River in South Africa, demonstrate that anatomical features that originated in *Homo sapiens* were present throughout the continent during this period. Additionally, archaeological sites associated with *Homo sapiens*, particularly from the Middle Stone Age, are widely distributed across Africa, including the site with the oldest footprint identified [2]. The question of whether these populations represent the direct ancestors of contemporary humans or represent isolated local populations remained unanswered. Due to recent advances in population genetic tools, more complex modelling and inference using larger datasets have become possible [3].

The discovery of Neanderthal admixture in Eurasian populations prompted research that suggested that an archaic hominin "ghost" population contributed to African populations [4]. These studies indicated that such a model could better explain the observed genetic data, particularly in western, southern, and central African populations. However, these studies contrasted only a single-origin model with an archaic hominin admixture model, leaving out other plausible models [1].

Challenging the Genetic Models: the study and its methodology

We aimed to discriminate between a broader set of demographic models by analysing the genomes of contemporary populations from diverse regions in Africa [1]. We considered four main models: single-population expansion, single-population expansion with regional persistence, archaic hominin admixture, and multi-regional evolution [5]. By including genetically and geographically diverse populations, we were able to infer demographic models that better explain the observed genetic diversity.

Critical to this study, we obtained whole-genome sequencing data for four African populations. The Nama from South Africa are an indigenous population that form part of a larger group of geographically close and culturally related individuals known collectively as the “Khoe-San”. The Khoe-San are reported to have the most divergent lineages of any other living population grouping and it is believed that they have largely remained isolated until ~2000 years ago. Participants from the Mende from Sierra Leone, the Gumuz from Ethiopia, and eastern African agriculturalists (Amhara and Oromo) were also included. British individuals were included as a representative of back-to-Africa gene flow and recent colonial admixture in South Africa. To account for gene flow from Neanderthals in Europeans, a high-coverage ancient Neanderthal genome was included in the analysis.

The results of our analyses confirmed the inadequacy of tree-like models and provided insights into more complex population structures. Rather than a simple single-origin model or archaic hominin admixture model, the best-fit models involved population reticulation, migration between early hominin populations, and divergence followed by merger events – a tangled vine with weakly separated offshoots instead of a tree of life (Figure 1) [6]. These “weakly structured stem” models indicated major stem lineages in southern, eastern, and

western/central Africa during the late Middle Pleistocene, followed by subsequent fragmentation and combination of subpopulations. The inferred models suggested that the Middle to Late Pleistocene was a critical period of change, with merger events between divergent stems likely influenced by shifts in wet and dry conditions across the African continent. The findings also highlighted the ongoing contribution of one of the stems to western Africans during the Last Glacial Maximum, indicating gene flow in western and/or central Africa. In addition to distinguishing between models of early human history, by including many present-day individuals and groups we can better predict and describe genetic variation for people living today, making these models more broadly useful.

It is often stated that “All models are wrong, but some are useful” (quote attributed to the statistician George Box). This highlights how constructing detailed models of human history is challenging, as model misspecification is inherent in such studies. It is also difficult to explore all plausible models, including the possibility of more complex models that involve additional stems or hybrid scenarios. The study's interpretations were also subject to uncertainties in estimating divergence times and migration rates, emphasizing the need for further research and testing with ancient DNA samples (though rarely found in Africa) and additional populations.

Why do we need to sequence additional African genomes?

Africa is home to the greatest level of genetic diversity in the world. Historically, genomic research has been biased towards populations of European and Asian ancestry, leading to an underrepresentation of African populations in genomic studies [7]. Results from these investigations may not be applicable to individuals with more diverse ancestries. To prevent further disparities, and to improve knowledge about human history, health and disease, it is imperative to include diverse genetic data sets. However, efforts have been made to address this imbalance and increase the representation of African genomes in genomic databases. Initiatives such as the African Genome Variation Project (AGVP) and the H3Africa (Human Heredity and Health in Africa) Consortium have been instrumental in collecting genomic data from diverse African populations [8,9]. A recent study which sequenced 180 individuals from 12 indigenous African populations identified millions of unreported variants, many with functional consequences [10]. Even so, there are definitely not enough African genomes sequenced (from modern or archaic samples) to represent the genetic diversity on the continent, but there is a drive by scientists to correct this [11].

The establishment of local sequencing facilities on the African continent is another important step, as some African countries do not allow the export of DNA samples. Locally in South Africa we are fortunate to have access to several sequencing facilities, such as the South African Medical Research Council Genomics Platform, as well as the Centre for Epidemic Response and Innovation (CERI) Genomics Centre at Stellenbosch University, to name but two [12]. While significant progress has been made in recent years, it is important to note that the sequencing of African genomes is an ongoing process, and there is still much work to be done. Access to African genomes will advance our knowledge of human evolution and improve our analyses of genomic variation linked to complex health traits.

How can African genomes contribute to health?

African genomes are not only key to deciphering human history, but can advance our understanding of how to use population-specific variants in health and disease so as to implement precision and preventative medicine on the continent. Many African countries face a significant burden from infectious diseases, as well as non-communicable diseases such as diabetes, cardiovascular disorders, and certain types of cancer. African genomes can aid in understanding the genetic basis of these diseases, identifying high-risk individuals, and developing targeted prevention and treatment strategies. This genetic data can also shed light on host-pathogen interactions and contribute to the development of vaccines and therapeutics.

However, to make precision and preventative medicine a reality for all, investigations of population genomics in diverse populations from across the world are crucial. As an example, a study involving South Africans that included Nama DNA as a reference population, identified ancestry-specific expression quantitative loci (eQTLs) associated with tuberculosis and type 2 diabetes that could potentially guide the search for new therapeutic targets for these diseases in African populations [13]. It is troubling that three African-specific eGenes would have been missed if we did not consider the genetic ancestry of the study participants. This was also the case in a genome-wide association study of tuberculosis [14].

It is unfortunate and concerning that the genetic data currently available are clearly not representative of the world's population. These datasets may also not be relevant to understudied groups and could even be unhelpful or misleading when determining genetic risk profiles for diseases in these settings, as a clear understanding of local population genomics is needed [15]. At the same time, developing economies, especially those in Africa, bear the brunt of socioeconomic inequalities, poor living conditions and disease. Primary healthcare incorporating genomics could assist developing economies in not only diagnosing and treating diseases, but by helping to prevent diseases, resulting in cost savings.

Societal impact

As mentioned above, a number of consortia have made significant contributions to include more diverse populations that have historically been underrepresented, and contributed to the training and development of African genetic researchers. Clearly global cooperation is key in these consortia, as our collaboration has also shown. We have worked together on a number of studies involving population genetics as well as genetic susceptibility to tuberculosis since 2009 [16–20]. This longstanding collaboration has resulted in many co-authored manuscripts, grants and the training of PhD students. However, none of this work would have been possible without the study participants. We have made several trips to provide feedback on the research, and build on community participation and collaboration. As part of this collaboration, local clinics were better equipped and community members were trained to assist with recruitment (Figure 2). Such efforts can help to promote science education and a culture of science in communities.

Conclusion

Our study highlights the critical role that DNA from contemporary Africans can play in understanding deep human history. By integrating the genetic data of a large number of contemporary individuals and groups, we can better anticipate and explain genetic variation in present-day individuals, allowing us apply these models to health-related research concerns. There is clearly still much to be learnt by focusing on genetic data from contemporary individuals, especially when ancient DNA – crucial in revealing intriguing history and answering important concerns - may not exist for the relevant time periods, as is typically the case in Africa. Overall, our study contributes to a better understanding of human and specifically African population history and highlights the limitations of simplistic models, encouraging the re-evaluation of previous interpretations of genomic and fossil data.

Acknowledgements

M.M. and E.H. acknowledge the support of the DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research, the South African Medical Research Council Centre for Tuberculosis Research, and the Division of Molecular Biology and Human Genetics at Stellenbosch University, Cape Town, South Africa.

Competing Interests

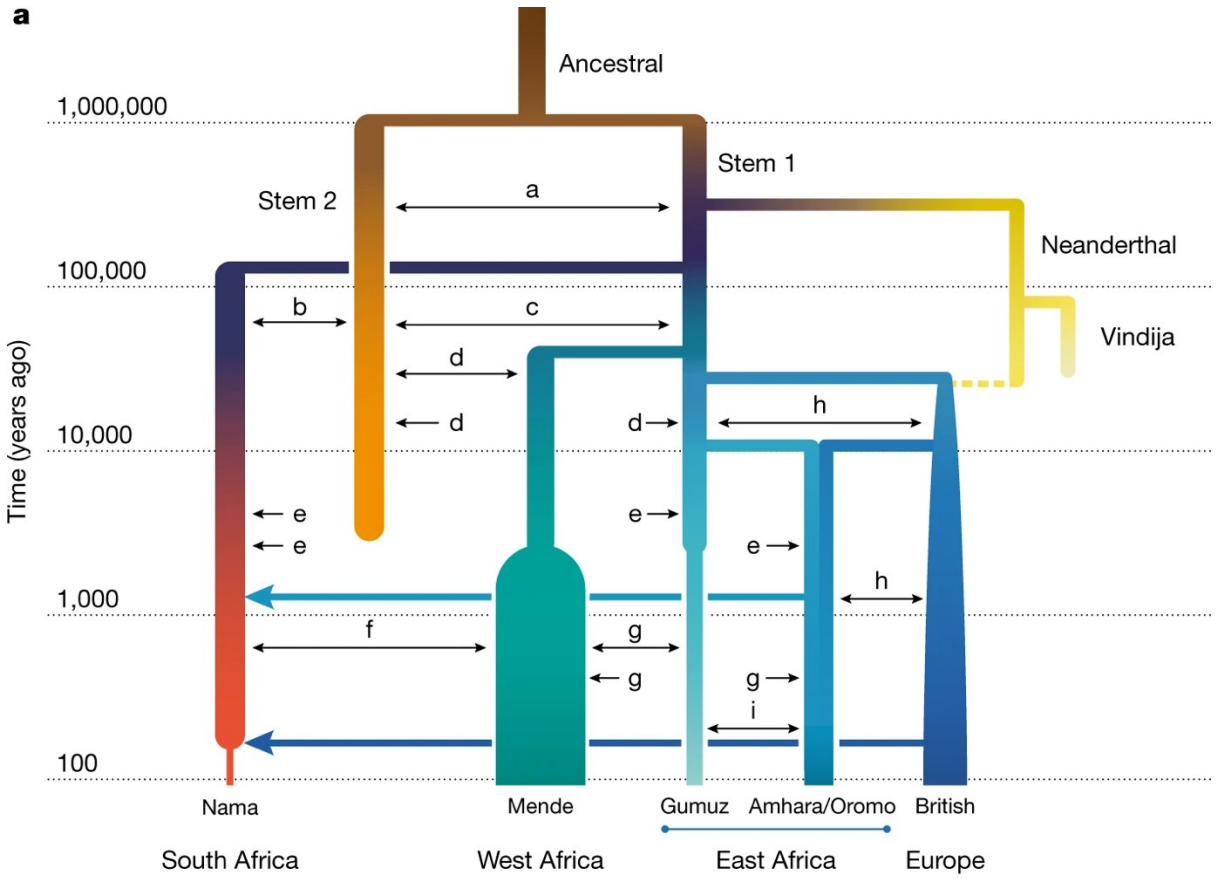
We have no competing interests to declare.

References

1. Ragsdale AP, Weaver TD, Atkinson EG, Hoal EG, Möller M, Henn BM, et al. A weakly structured stem for human origins in Africa. Nature. 2023;617(7962):755–63. https://doi.org/10.1038/s41586-023-06055-y.
2. Helm CW, Carr AS, Lockley MG, Cawthra HC, De Vynck JC, Dixon MG, et al. Dating the Pleistocene hominin ichnosites on South Africa's Cape south coast. Ichnos. 2023;0(0):1–20. https://doi.org/10.1080/10420940.2023.2204231.
3. Scerri EML, Thomas MG, Manica A, Gunz P, Stock JT, Stringer C, et al. Did Our Species Evolve in Subdivided Populations across Africa, and Why Does It Matter? Trends Ecol Evol. 2018;33(8):582–94. https://doi.org/10.1016/j.tree.2018.05.005.
4. Hammer MF, Woerner AE, Mendez FL, Watkins JC, Wall JD. Genetic evidence for archaic admixture in Africa. Proc Natl Acad Sci U S A. 2011;108(37):15123–8. https://doi.org/10.1073/pnas.1109300108.
5. Henn BM, Steele TE, Weaver TD. Clarifying distinct models of modern human origins in Africa. Curr Opin Genet Dev. 2018;53:148–56. https://doi.org/10.1016/j.gde.2018.10.003.
6. Scerri EML. One species, many roots? Nat Ecol Evol. 2023:1–2. https://doi.org/10.1038/s41559-023-02080-2.
7. Martin AR, Teferra S, Möller M, Hoal EG, Daly MJ. The critical needs and challenges for genetic architecture studies in Africa. Curr Opin Genet Dev. 2018;53:113–20. https://doi.org/10.1016/j.gde.2018.08.005.
8. Choudhury A, Aron S, Botigué LR, Sengupta D, Botha G, Bensellak T, et al. High-depth African genomes inform human migration and health. Nature. 2020;586(7831):741–8. https://doi.org/10.1038/s41586-020-2859-7.
9. Gurdasani D, Carstensen T, Tekola-Ayele F, Pagani L, Tachmazidou I, Hatzikotoulas K, et al. The African Genome Variation Project shapes medical genetics in Africa. Nature. 2015;517(7534):327–32. https://doi.org/10.1038/nature13997.
10. Fan S, Spence JP, Feng Y, Hansen MEB, Terhorst J, Beltrame MH, et al. Whole-genome sequencing reveals a complex African population demographic history and signatures of local adaptation. Cell. 2023;186(5):923-939.e14. https://doi.org/10.1016/j.cell.2023.01.042.
11. Wonkam A. Sequence three million genomes across Africa. Nature. 2021;590(7845):209–11. https://doi.org/10.1038/d41586-021-00313-7.
12. Glanzmann B, Jooste T, Ghoor S, Gordon R, Mia R, Mao J, et al. Human whole genome sequencing in South Africa. Sci Rep. 2021;11(1):606. https://doi.org/10.1038/s41598-020-79794-x.
13. Swart Y, Uren C, Eckold C, Cliff JM, Malherbe ST, Ronacher K, et al. cis-eQTL mapping of TB-T2D comorbidity elucidates the involvement of African ancestry in TB susceptibility. bioRxiv. 2022:2022.10.19.512814. https://doi.org/10.1101/2022.10.19.512814.
14. Swart Y, Uren C, van Helden PD, Hoal EG, Möller M. Local Ancestry Adjusted Allelic Association Analysis Robustly Captures Tuberculosis Susceptibility Loci. Front Genet. 2021;12:716558. https://doi.org/10.3389/fgene.2021.716558.
15. Martin AR, Teferra S, Möller M, Hoal EG, Daly MJ. The critical needs and challenges for genetic architecture studies in Africa. Curr Opin Genet Dev. 2018;53:113–20. https://doi.org/10.1016/j.gde.2018.08.005.
16. Smith MH, Myrick JW, Oyageshio O, Uren C, Saayman J, Boolay S, et al. Epidemiological correlates of overweight and obesity in the Northern Cape Province, South Africa. PeerJ. 2023;11:e14723. https://doi.org/10.7717/peerj.14723.

17. Reynolds AW, Grote MN, Myrick JW, Al-Hindi DR, Siford RL, Mastoras M, et al. Persistence of Matrilocal Postmarital Residence Across Multiple Generations in Southern Africa. Hum Nat. 2023. https://doi.org/10.1007/s12110-023-09452-4.
18. Schurz H, Kinnear CJ, Gignoux C, Wojcik G, van Helden PD, Tromp G, et al. A Sex-Stratified Genome-Wide Association Study of Tuberculosis Using a Multi-Ethnic Genotyping Array. Front Genet. 2018;9:678. https://doi.org/10.3389/fgene.2018.00678.
19. van Eeden G, Uren C, Pless E, Mastoras M, van der Spuy GD, Tromp G, et al. The recombination landscape of the Khoe-San likely represents the upper limits of recombination divergence in humans. Genome Biol. 2022;23(1):172. https://doi.org/10.1186/s13059-022-02744-5.
20. Uren C, Kim M, Martin AR, Bobo D, Gignoux CR, van Helden PD, et al. Fine-Scale Human Population Structure in Southern Africa Reflects Ecogeographic Boundaries. Genetics. 2016;204(1):303–14. https://doi.org/10.1534/genetics.116.187369.

Figures


a
Ancestral
1,000,000
Stem 1
Stem 2
a
Neanderthal
100,000
b
c
Vindija
d
d
d
h
10,000
e
e
e
e
h
1,000
f
g
g
g
i
100
Time (years ago)
Nama
Mende
Gumuz
Amhara/Oromo
British
South Africa
West Africa
East Africa
Europe

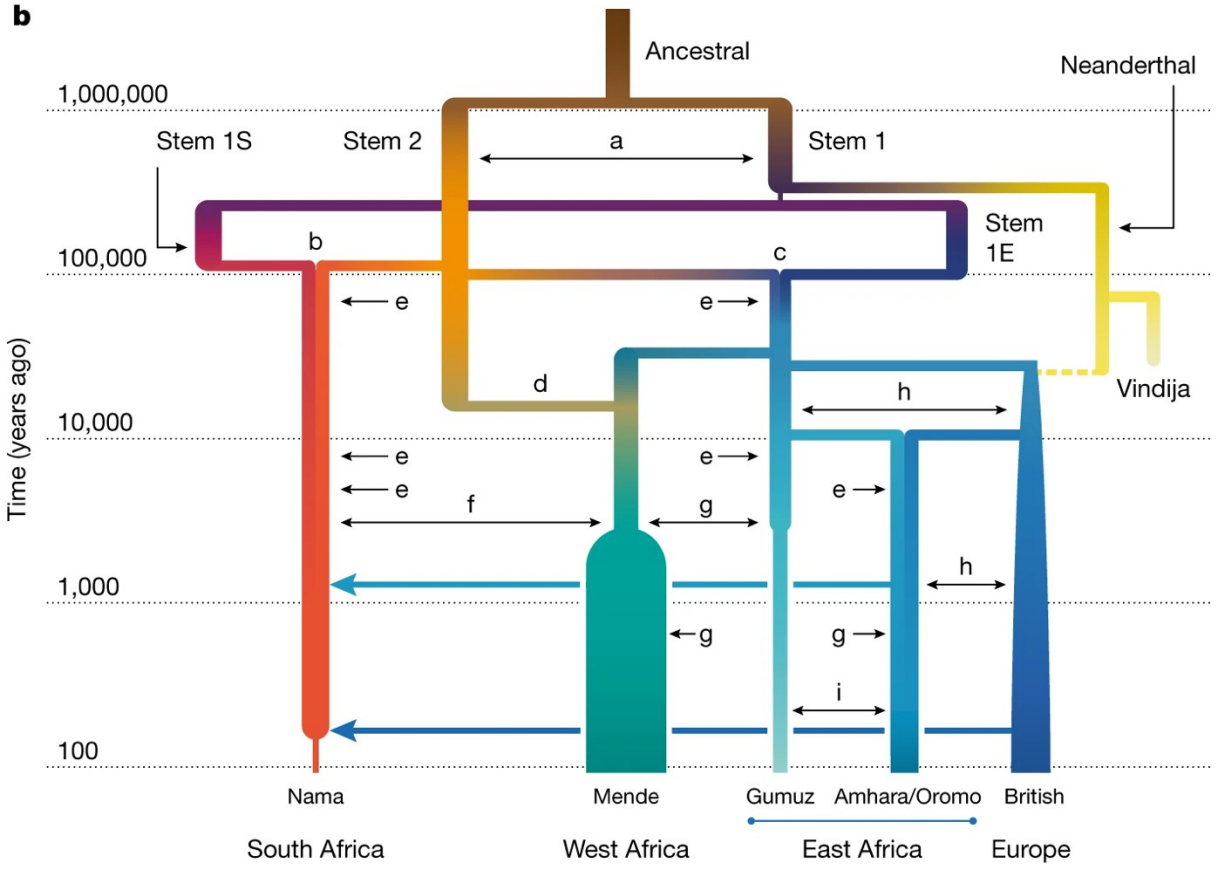

b
Ancestral
Neanderthal
1,000,000
Stem 1S
Stem 2
a
Stem 1
Stem
1E
100,000
b
c
e
e
d
h
Vindija
10,000
e
e
e
e
f
g
h
1,000
g
g
i
100
Time (years ago)
Nama
Mende
Gumuz
Amhara/Oromo
British
South Africa
West Africa
East Africa
Europe

Figure 1: New “weakly structured stem” models of human evolution with a) continuous migration and b) multiple mergers: where stem populations separated, but continually exchanged genetic material. From Ragsdale et al [1] and licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

Figure 2: Human genetics research is built on community participation and collaboration. As part of our research, local clinics were better equipped and community members were trained to assist with recruitment. Pictured here are study co-investigators, the study nurse and community members involved in recruitment.