

How the language you speak aligns to your genetic origins and may impact research on your health

April 20 2021

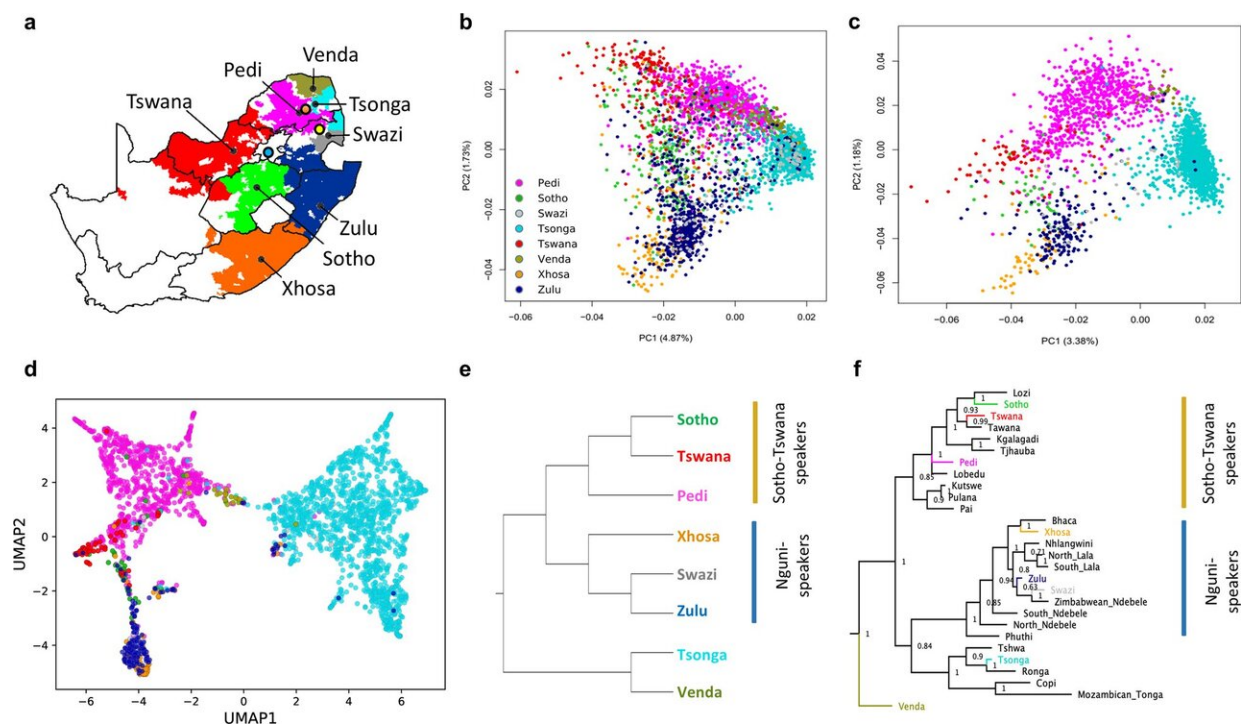


Fig. 1: Population structure and genetic affinities of South-Eastern Bantu-speaking (SEB) groups from South Africa correspond to both linguistic phylogeny and geographic distribution. Credit: *Nature Communications* (2021). DOI: 10.1038/s41467-021-22207-y

A new study challenges the presumption that all South-Eastern-Bantu

speaking groups are a single genetic entity.

The South-Eastern-Bantu (SEB) language family includes isiZulu, isiXhosa, siSwati, Xitsonga, Tshivenda, Sepedi, Sesotho and Setswana.

Almost 80% of South Africans speak one of the SEB family languages as their first language. Their origins can be traced to farmers of West-Central Africa whose descendants over the past two millennia spread south of the equator and finally into Southern Africa.

Since then, varying degrees of sedentism [the practice of living in one place for a long time], population movements and interaction with Khoe and San communities, as well as people speaking other SEB languages, ultimately generated what are today distinct Southern African languages such as isiZulu, isiXhosa and Sesotho.

Despite these linguistic differences, these groups are treated mostly as a single group in genetic studies.

Understanding [genetic diversity](#) in a population is critical to the success of disease genetic studies. If two genetically distinct populations are treated as one, the methods normally used to find disease genes could become error prone.

Consideration of these genetic differences is critical to providing a reliable understanding of the genetics of complex diseases, such as diabetes and hypertension, in South Africans.

Dr. Dhriti Sengupta and Dr. Ananyo Choudhury in the Sydney Brenner Institute for Molecular Bioscience (SBIMB) at Wits University were joint lead authors of the paper published in *Nature Communications* on 7 April 2021.

The study comprised a multidisciplinary team of geneticists, bioinformaticians, linguists, historians and archaeologists from Wits University (Michèle Ramsay, Scott Hazelhurst, Shaun Aron and Gavin Whitelaw), the University of Limpopo, and partners in Belgium, Sweden and Switzerland.

"South Eastern Bantu-speakers have a clear linguistic division—they speak more than nine distinct languages—and their geography is clear: some of the groups are found more frequently in the north, some in central, and some in southern Africa. Yet despite these characteristics, the SEB groups have so far been treated as a single genetic entity," says Choudhury.

The study found that SEB speaking groups are too different to be treated as a single genetic unit.

"So if you are treating say, Tsonga and Xhosa, as the same population—as was often done until now—you might get a completely wrong gene implicated for a disease," says Sengupta.

About the study

The study, titled: Genetic substructure and complex demographic history of South African Bantu speakers aimed to find out whether the SEB speakers are indeed a single genetic entity or if they have enough genetic differences to be grouped into smaller units.

Genetic data from more than 5000 participants speaking eight different southern African languages were generated and analyzed.

These languages are isiZulu, isiXhosa, siSwati, Xitsonga, Tshivenda, Sepedi, Sesotho and Setswana.

Participants were recruited from research sites in Soweto in Gauteng, Agincourt in Mpumalanga, and Dikgale in Limpopo province.

Genetic differences reflect geography, language and history

The study detected major variations in genetic contribution from the Khoe and San into SEB speaking groups; some groups have received a lot of genetic influx from Khoe and San people, while others have had a very little genetic exchange with these groups.

This variation ranged on average from about 2% in Tsonga to more than 20% in Xhosa and Tswana.

This suggests that SEB speaking groups are too different to be treated as a single genetic unit.

"The study showed that there could be substantial errors in disease gene discovery and disease risk estimation if the differences between South-Eastern-Bantu speaking groups are not taken into consideration," says Sengupta.

The [genetic data](#) also show major differences in the history of these groups over the last 1000 years. Genetic exchanges were found to have occurred at different points in time, suggesting a unique journey of each group across the southern African landscape over the past millennium.

These genetic differences are strong enough to impact the outcomes of biomedical genetic research.

Sengupta emphasizes, however, that ethnolinguistic identities are complex and cautioned against extrapolating broad conclusions from the

findings regarding [genetic differences](#).

"Although genetic data showed differences [separation] between groups, there was also a substantial amount of overlap [similarity]. So while findings regarding differences could have huge value from a research perspective, they should not be generalized," she says.

A genetic blueprint for future health

A common approach to identify if a genetic variant causes or predisposes us to a disease is to take a set of individuals with a disease (e.g., high blood pressure or diabetes) and another set of healthy individuals without the disease, and then compare the occurrence of many genetic variants in the two sets.

If a variant shows a notable frequency difference between the two sets it is assumed that the genetic variant could be associated with the disease.

"However, this approach depends entirely on the underlying assumption that the two groups consist of genetically similar individuals. One of the major highlights of our study is the observation that Bantu-speakers from two geographic regions—or two ethnolinguistic groups—cannot be treated as if they are the same when it comes to [disease genetic studies](#)," says Choudhury.

Future studies, especially those testing a small number of variants, need to be more nuanced and have balanced ethnolinguistic and geographic representation, he says.

This study is the second landmark study in African population genetics, published in the last six months, led by researchers in the Sydney Brenner Institute for Molecular Bioscience in the Faculty of Health Sciences at Wits University.

Professor Michèle Ramsay, director of the SBIMB and corresponding author of the study, says: "The in-depth analysis of several large African genetic datasets has just begun. We look forward to mining these datasets to provide new insights into key population histories and the genetics of complex diseases in Africa."

More information: Genetic substructure and complex demographic history of South African Bantu speakers, *Nature Communications* (2021). [DOI: 10.1038/s41467-021-22207-y](https://doi.org/10.1038/s41467-021-22207-y)

Provided by Wits University

Citation: How the language you speak aligns to your genetic origins and may impact research on your health (2021, April 20) retrieved 2 May 2024 from <https://phys.org/news/2021-04-language-aligns-genetic-impact-health.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.