



# African Genome Variation Project Enlightens Population History, Provides Basis for Disease Studies

Dec 03, 2014 | [Julia Karow](#)

## *Premium*

NEW YORK (GenomeWeb) – A new study on genetic diversity in Africa, called the African Genome Variation Project, shines light on the population history of sub-Saharan Africa and provides the basis for future medical genetics studies, including a new pan-African genotype array.

The project, [published online in \*Nature\* today](#), analyzed the genomes of about 1,800 individuals from 18 ethno-linguistic groups in Western, Eastern, and Southern Africa, most by array genotyping and some by whole-genome sequencing. It was conducted as a collaboration between the Wellcome Trust Sanger Institute in the UK, the US National Institutes of Health, the UK Medical Research Council, and research institutions in South Africa, Ethiopia, the Gambia, Ghana, and elsewhere.

According to the analysis, many sub-Saharan African populations have genetic contributions from Eurasians and hunter-gatherers, providing clues to ancient population movements.

The study also identified new disease susceptibility loci that have been under positive selection, including genes associated with hypertension, malaria, and other infectious diseases.

Based on their results and additional data, the scientists are proposing to design a pan-African genotype array that can effectively capture common genetic variation across Africa. Together with improved imputation panels that can infer missing genotypes, such arrays will be useful for future medical genomic studies in Africa, they said.

According to Manjinder Sandhu, one of the senior authors of the study and a researcher at the Sanger Institute and the University of Cambridge, the motivation for the project was to conduct large-scale medical genetics studies in Africa, which require a better map of the genetic diversity on the continent. Prior to their study, "the relevant genetic catalog to describe individual variation across Africa to create the best studies did not exist," he told GenomeWeb, prompting him and his colleagues to start collecting samples from various populations for analysis.

"The idea was to choose populations that represented ongoing medical genetics studies in Africa," he

said, including both large population groups as well as smaller hunter-gatherer populations that tend to be more genetically divergent.

Most of the samples were collected by members of the African Partnership for Chronic Disease Research, APCDR, which overlap in part with the pan-African Human Heredity and Health in Africa, H3Africa, initiative, he said.

For the project, the Sanger scientists analyzed 1,481 individuals using the Illumina HumanOmni2.5M genotyping array and sequenced the genomes of 320 others. For their analysis, they also added publicly available data from the 1,000 Genomes Project.

Genetic diversity was substantial – from the sequence data, the researchers identified about 30 million SNPs, up to a quarter of which were novel and up to another quarter were unshared between populations.

Although African populations are generally more diverse than populations outside of the continent, genetic diversity among African populations was not as great as previously assumed, Sandhu said. This likely reflects population expansions and movements, for example, the so-called Bantu expansion that started in West Africa about 3,000 to 5,000 years ago.

He and his colleagues also found regionally distinct and complex admixture with other populations. For example, several African populations showed genetic contributions from Eurasians, which might stem from Eurasians migrating back to Africa up to 10,500 years ago. In addition, some populations had genetic ancestry from ancient hunter-gatherers that may have joined Bantu populations at different points in time.

In terms of medical genetics, the researchers found both known and new disease susceptibility loci that appear to have undergone positive selection in response to the local environment, including gene regions associated with malaria, hypertension, sickle cell anemia, Lassa fever, and trypanosomiasis.

Based on their results, the researchers assessed how future medical genetics studies in Africa would be best designed. According to Sandhu, such studies should combine diverse whole-genome sequence reference panels for imputation that broadly represent populations across Africa with a new design for a pan-African genotype array. Genetic association studies require large-scale studies, he said, which are currently best served by relatively low-cost arrays coupled with imputation, rather than by whole-genome sequencing.

The reason why a new array design is needed is that current chips are largely based on European populations that do not capture the genetic diversity of Africa, said Deepti Gurdasani, the lead author of the study and a postdoctoral fellow at the Sanger Institute.

Using a potential chip design with just one million variants, they were able to capture more than 80 percent of common genetic variation in Africans, which she said is encouraging, "because people have often thought that you need many more variants to capture diversity across Africa."

To improve the reference panel of African genomes, the researchers have already collected almost 3,000 whole African genomes sourced from public resources, other investigators, and additional

genomes sequenced at the Sanger Institute. They are hoping to compile a reference panel of at least 4,000 whole genomes by the middle of next year, Sandhu said, adding that a key outcome of that panel will be to design a new genotype array for populations across Africa.

Sandhu said the researchers are currently in discussions with two potential chip manufacturers, Illumina and Affymetrix, and are hoping to have a new chip available sometime next year, which would be used for both ongoing and new medical genetics studies.

Current medical genetics studies with African partners, which are primarily funded by the NIH, the Wellcome Trust, and the UK's Medical Research Council, are focused on diseases and conditions such as diabetes, hypertension, blood cholesterol, and liver function, Sandhu said, with others planned to look into response to vaccination and chronic infection with HIV and hepatitis C.

The results could serve not only Africans but also provide a better understanding of global disease susceptibility, he said, allowing researchers to identify causal variants by using an approach called trans-ethnic fine-mapping.

Filed Under [Genetic Research](#) [Sequencing Technology](#) [Microarrays & Multiplexing](#)

[hypertension](#) [malaria](#) [Sanger Institute](#) [genotyping array](#) [whole-genome sequencing](#)



**[Get Weekly Sequencing Technology Updates](#)**



**[Get Weekly Microarrays & Multiplexing Updates](#)**

## Related Articles

---

Sep 17, 2014

**[Ancient Genome Analysis Indicates Three Ancestral Populations Contributed to Modern-day Europeans](#)**

Dec 04, 2014

**[Study Suggests African Khoisan Hunter-Gatherers Remained Isolated, Represented Largest Population](#)**

Mar 28, 2014

**[Using Genotyping Tools, Researchers Trace Cattle Domestication, Breed History](#)**

Mar 13, 2014

**[Lactase Persistence Study Offers Look at Past Migration, Pastoralism in Africa](#)**

---

Apr 03, 2014

**Genomics in the Journals**

---

Jan 27, 2015

**Illumina, Collaborators Design Multi-Ethnic Genotyping Array to Empower GWAS in Diverse Populations**

---

[Privacy Policy](#). Copyright © 2015 Genomeweb LLC. All Rights Reserved.