

Newly sequenced genome reveals coffee's prehistoric origin story, and its future under climate change

April 15 2024



University at Buffalo researchers have created what they say is the highest quality reference genome to date of the world's most popular coffee species, Arabica. Credit: University at Buffalo

The key to growing coffee plants that can better resist climate change in the decades to come may lie in the ancient past.

Researchers co-led by the University at Buffalo have created what they say is the highest quality [reference genome](#) to date of the world's most popular coffee species, Arabica, unearthing secrets about its lineage that span millennia and continents.

Their [findings](#), published in *Nature Genetics*, suggest that *Coffea arabica* developed more than 600,000 years ago in the forests of Ethiopia via natural mating between two other coffee species. Arabica's population waxed and waned throughout Earth's heating and cooling periods over thousands of years, the study found, before eventually being cultivated in Ethiopia and Yemen, and then spread over the globe.

"We've used genomic information in plants alive today to go back in time and paint the most accurate picture possible of Arabica's long history, as well as determine how modern cultivated [varieties](#) are related to each other," says the study's co-corresponding author, Victor Albert, Ph.D., Empire Innovation Professor in the UB Department of Biological Sciences, within the College of Arts and Sciences.

Coffee giants like Starbucks and Tim Hortons exclusively use beans from Arabica plants to brew the millions of cups of coffee they serve everyday, yet, in part due to a low genetic diversity stemming from a history of inbreeding and small population size, Arabica is susceptible to many pests and diseases and can only be cultivated in a few places in the world where pathogen threats are lower and climate conditions are more favorable.

"A detailed understanding of the origins and breeding history of contemporary varieties are crucial to developing new Arabica cultivars better adapted to [climate change](#)," Albert says.

From their new reference genome, accomplished using cutting-edge DNA sequencing technology and advanced data science, the team was able to sequence 39 Arabica varieties and even an 18th century specimen used by Swedish naturalist Carl Linnaeus to name the species.

The reference genome is now available in a publicly available digital database.

"While other public references for Arabica coffee do exist, the quality of our team's work is extremely high," says one of the study's co-leaders, Patrick Descombes, senior expert in genomics at Nestlé Research. "We used state-of-the-art genomics approaches—including long- and short-read high throughput DNA sequencing—to create the most advanced, complete and continuous Arabica reference genome to date."

Humanity's favorite coffee evolved without people's help

Arabica is the source of approximately 60% of the world's total coffee products, with its seeds helping millions start their day or stay up late. However, the initial crossbreeding that created it was done without any intervention from humans.

Arabica formed as a natural hybridization between *Coffea canephora* and *Coffea eugenioides*, whereupon it received two sets of chromosomes from each parent. Scientists have had a hard time pinpointing exactly when—and where—this allopolyploidization event took place, with estimates ranging everywhere from 10,000 to 1 million years ago.

To find evidence for the original event, UB researchers and their partners ran their various Arabica genomes through a computational modeling program to look for signatures of the species' foundation.

The models show three population bottlenecks during Arabica's history, with the oldest happening some 29,000 generations—or 610,000 years—ago. This suggests Arabica formed sometime before that, anywhere from 610,000 to 1 million years ago, researchers say.

"In other words, the crossbreeding that created Arabica wasn't something that humans did," Albert says. "It's pretty clear that this polyploidy event predated modern humans and the cultivation of coffee."

Coffee plants have long been thought to have developed in Ethiopia, but varieties that the team collected around the Great Rift Valley, which stretches from Southeast Africa to Asia, displayed a clear geographic split. The wild varieties studied all originated from the western side, while the cultivated varieties all originated from the eastern side closest to the Bab al-Mandab strait that separates Africa and Yemen.

That would align with evidence that coffee cultivation may have started principally in Yemen, around the 15th century. Indian monk Baba Budan is believed to have smuggled the fabled "seven seeds" out of Yemen around 1600, establishing Indian Arabica cultivars and setting the stage for coffee's global reach today.

"It looks like Yemeni coffee diversity may be the founder of all of the current major varieties," Descombes says. "Coffee is not a crop that has been heavily crossbred, such as maize or wheat, to create new varieties. People mainly chose a variety they liked and then grew it. So the varieties we have today have probably been around for a long time."

How climate impacted Arabica's population

East Africa's geoclimatic history is well documented due to research on human origins, so researchers could contrast climate events with how the wild and cultivated Arabica populations fluctuated over time.

Modeling shows a long period of low population size between 20–100,000 years ago, which roughly coincides with an extended drought and cooler climate believed to have hit the region between 40–70,000 years ago. The population then increased during the African humid period, around 6–15,000 years ago, when growth conditions were likely more beneficial.

During this same time, around 30,000 years ago, the wild varieties and the varieties that would eventually become cultivated by humans split from each other.

"They still occasionally bred with each other, but likely stopped around the end of the African humid period and the widening of the strait due to rising sea levels around 8,000 to 9,000 years ago," says Jarkko Salojärvi, assistant professor at Nanyang Technological University in Singapore and another co-corresponding author of work.

Low genetic diversity threatens Arabica

Cultivated Arabica is estimated to have an effective population size of only 10,000 to 50,000 individuals. Its low genetic diversity means it could be completely decimated, like the monoculture Cavendish banana, by pathogens, such as coffee leaf rust, which causes \$1–2 billion in losses annually.

The reference genome was able to shed more light on how one line of Arabica varieties obtained strong resistance to the disease.

The Timor variety formed in Southeast Asia as a spontaneous hybrid between Arabica and one of its parents, *Coffea canephora*. Also known as Robusta and used primarily for instant coffee, this species is more resistant to disease than Arabica.

"Thus, when Robusta hybridized itself back into Arabica on Timor, it brought some of its pathogen defense genes along with it," says Albert, who also co-led [sequencing of the Robusta genome in 2014](#). Albert and collaborators' current work also presents a highly improved version of the Robusta genome, as well as new sequence of Arabica's other progenitor species, *Coffea eugenioides*.

While breeders have tried replicating this crossbreeding to boost pathogen defense, the new Arabica reference genome allowed the present researchers to pinpoint a novel region harboring members of the RPP8 resistance gene family as well as a general regulator of resistance genes, CPR1.

"These results suggest a novel target locus for potentially improving pathogen resistance in Arabica," Salojärvi says.

The genome provided other new findings as well, like which wild varieties are closest to modern, cultivated Arabica coffee. They also found that the Typica variety, an early Dutch cultivar originating from either India or Sri Lanka, is likely the parent of the Bourbon variety, principally cultivated by the French.

"Our work has not been unlike reconstructing the family tree of a very important family," Albert says.

More information: Jarkko Salojärvi et al, The genome and population genomics of allopolyploid *Coffea arabica* reveal the diversification history of modern coffee cultivars, *Nature Genetics* (2024). [DOI: 10.1038/s41588-024-01695-w](https://doi.org/10.1038/s41588-024-01695-w)

Provided by University at Buffalo

Citation: Newly sequenced genome reveals coffee's prehistoric origin story, and its future under climate change (2024, April 15) retrieved 2 May 2024 from

<https://phys.org/news/2024-04-newly-sequenced-genome-reveals-coffee.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.