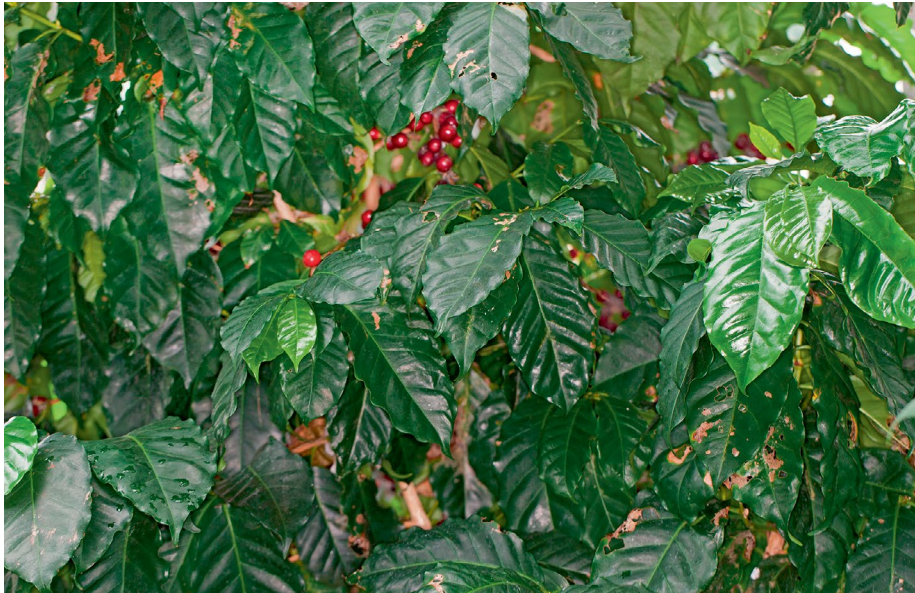


## Research highlights

Coffee genomics

<https://doi.org/10.1038/s41477-024-01714-y>

# Coffee history under the genomic lens



Arabian coffee (*Coffea arabica*) cherries on a twig.

In the global market for hot drinks, coffee holds a unique important position. About 60% of coffee products come from the species *Coffea arabica* (Arabica coffee, CCEE), which is a spontaneous allotetraploid hybrid of *Coffea eugenioides* (EE) and *Coffea canephora* (Robusta coffee, CC). Previous research and historical documentation gave us useful but preliminary information about the evolution of *Coffea arabica*. To deepen our understanding, Jarkko Salojärvi, from the Nanyang Technological University, Singapore, and colleagues integrated up-to-date genome technologies with population genetics analyses to dissect the evolution history of modern coffee in detail.

The genomes of Arabica coffee and its two progenitor representatives were sequenced and assembled using a hybrid assembly strategy. Pacbio HiFi technology was used to further improve the Arabica genome,

yielding a highly continuous assembly that showed no sign of long-terminal repeats transposable element mobilization but two harmoniously coexisting subgenomes. The choice of a di-haploid Arabica, as opposed to a normal tetraploid accession, was an important element to generate a highly continuous and advanced genome assembly. The two subgenomes showed high conservation with the modern representatives of the two diploid progenitors and between each other. No overall subgenome dominance was observed at global gene expression level, but rather a rich mosaic-like pattern where the expression bias varied according to the gene family, indicating an early stage of diploidization.

The researchers then re-sequenced 46 *Coffea* accessions including 41 *C. arabica*. Occasional homoeologous exchanges between the subgenomes were observed

to contribute to *C. arabica* diversity, and possibly also played a role in the establishment of a stable allopolyploid line after the hybridization event. The wild population has an already low diversity, mostly caused by the allopolyploid founding event that occurred 610–350 thousand years ago (ka). Analyses with wild Arabica coffee showed a prolonged period of small population sizes after the hybridization event and a more recent bottleneck around 5 ka. Compared to the wild, only slight nucleotide diversity reduction was observed in the cultivars, so domestication may have caused only minor bottleneck.

Further analyses showed kinship relationships among the cultivars and wild accessions, and inferred an admixture event shared by the cultivars and some of the wild individuals; this heterozygosity-increasing event may have contributed towards hybrid vigor effects in the cultivated coffee. The analyses further pointed to the Gesha region as a hotspot of wild accessions amenable to domestication.

Cultivars and wild accessions are highly differentiated in regions enriched for defense related genes, so domestication possibly reduced the diversity of immune defense genes in cultivars. The five sequenced Timor hybrid lines all shared an introgression from Robusta coffee on chromosome 4, which contained resistance-related genes and likely underpinned their high resistance to coffee leaf rust. Strong candidate genes for leaf rust resistance were found by overlapping genes in selected regions with those displaying expression change in response to leaf rust inoculation, which may be useful for coffee breeding.

**Jun Lyu**

Nature Plants

Original reference: *Nat. Genet.* **56**, 721–731 (2024)