

Tea tree genome contains clues about how one leaf produces so many flavors

May 1 2017



A close up of a *Camellia sinensis* shrub. Credit: LiZhi Gao Lab

The most popular varieties of tea—including black tea, green tea,

Oolong tea, white tea, and chai—all come from the leaves of the evergreen shrub *Camellia sinensis*, otherwise known as the tea tree. Despite tea's immense cultural and economic significance, relatively little is known about the shrub behind the tea leaves. However, the first draft of the tea tree genome published May 1 in the journal *Molecular Plant* may help explain why tea leaves are so rich in antioxidants and caffeine.

Understanding how the tea tree genetically differs from its close relatives may help tea growers figure out what makes *Camellia sinensis* leaves so special. The genus *Camellia* contains over 100 species—including several popular decorative garden plants and *C. oleifera*, which produces "tea tree" oil—but only two major varieties (*C. sinensis* var. *assamica* and *C. sinensis* var. *sinensis*) are grown commercially for making tea. "There are many diverse flavors, but the mystery is what determines or what is the genetic basis of tea flavors?" says plant geneticist Lizhi Gao of Kunming Institute of Botany in China.

Previous studies have suggested that tea owes much of its flavor to a group of antioxidants called flavonoids, molecules that are thought to help plants survive in their environments. One, a bitter-tasting flavonoid called catechin, is particularly associated with tea flavor. Levels of catechin and other flavonoids vary among *Camellia* species, as does caffeine. Gao and his colleagues found that *C. sinensis* leaves not only contain high levels of catechins, caffeine, and flavonoids, but also have multiple copies of the [genes](#) that produce caffeine and flavonoids.



Researchers from Kunming Institute of Botany gathering tea tree leaves. Credit: Chao Shi

Caffeine and flavonoids such as catechins are not proteins (and therefore not encoded in the genome directly), but genetically encoded proteins in the tea leaves manufacture them. All *Camellia* species have genes for the caffeine- and flavonoid-producing pathways, but each species expresses those genes at different levels. That variation may explain why *C. sinensis* leaves are suitable for making tea, while other *Camellia* species' leaves aren't.

Gao and his colleagues estimate that more than half of the base pairs (67%) in the tea tree genome are part of retrotransposon sequences, or

"jumping genes", which have copied-and-pasted themselves into different spots in the genome numerous times. The large number of retrotransposons resulted in a dramatic expansion in genome size of tea tree, and possibly many, many duplicates of certain genes, including the disease-resistant ones. The researchers think that these "expanded" gene families must have helped tea [trees](#) adapt to different climates and environmental stresses, as tea trees grow well on several continents in a wide range of climate conditions. Since much of the retrotransposon copying & pasting seems to have happened relatively recently in the tea tree's evolutionary history, the researchers theorize that at least some of the duplications are responses to cultivation.

However, these duplicated genes and the large number of repeat sequences also turned assembling a tea tree genome into an uphill battle. "Our lab has successfully sequenced and assembled more than twenty plant genomes," says Gao. "But this genome, the tea tree genome, was tough."



Li-zhi Gao collecting tea tree leaves to be sequenced. Credit: Yong-sheng Yi

For one thing, the tea tree genome turned out to be much larger than initially expected. At 3.02 billion base pairs in length, the tea tree genome is more than four times the size of the coffee plant genome and much larger than most sequenced plant species. Further complicating the picture is the fact that many of those genes are duplicates or near-duplicates. Whole genomes are too long to sequence in one piece, so instead, scientists must copy thousands upon thousands of genome fragments, sequence them, and identify overlapping sequences that appear in multiple fragments. Those overlap sites become sign posts for lining up the fragments in the correct order. However, when the genome itself contains sequences that are repeated hundreds or thousands of times, those overlaps disappear into the crowd of repeats; it's like assembling a million piece puzzle where all the middle pieces look almost exactly alike.

All told, even with modern sequencing, assembling the genome took the team over 5 years.

And still, there is more work to do, both in terms of double-checking the genome draft and in terms of sequencing different tea tree varieties from around the world. "Together with the construction of genetic maps and new sequencing technologies, we are working on an updated tea tree [genome](#) that will investigate some of the flavor," says Gao. "We will look at gene copy number variation to see how they affect tea properties, like flavor. We want to get a map of different tea tree variation and answer how it was domesticated, cultivated, and dispersed to different continents of the world."

More information: *Molecular Plant*, Xia et al.: "The tea tree genome provides insights into tea flavor and independent evolution of caffeine biosynthesis" [www.cell.com/molecular-plant/f ...](http://www.cell.com/molecular-plant/fulltext/S1674-2052(17)30103-X)
[1674-2052\(17\)30103-X](https://doi.org/10.1016/j.molp.2017.04.002) , DOI: [10.1016/j.molp.2017.04.002](https://doi.org/10.1016/j.molp.2017.04.002)

Provided by Cell Press

Citation: Tea tree genome contains clues about how one leaf produces so many flavors (2017, May 1) retrieved 20 September 2024 from <https://phys.org/news/2017-05-tea-tree-genome-clues-leaf.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.