

Sequencing the watermelon family tree reveals ‘lost’ disease-resistance genes that were bred out generations ago



s autumn looms, we’re enjoying the last bites of sweet, juicy watermelon.

Conventional agriculture has molded our fruits and veggies to suit our palates, gradually crafting domesticated *Citrullus lanatus* from three ancestral melon species. But the process may have also removed valuable traits.

Researchers at the Boyce Thompson Institute in Ithaca, New York, have analyzed genomes of watermelon and its ancestors, revealing traits that early breeders may have inadvertently removed in their quest to maximize the red, sweet, watery flesh of the fruit. Their report appears in [Plant Biotechnology Journal](#).

“Pangenomics” is a newish term, referring to ways that a species’ genome can vary, DNA base by DNA base. I wrote The Age of the Pangenome Dawns here at [DNA Science](#) last year, about the [Human Pangenome Reference Consortium](#). Their goal is to “capture all human genome variation and support research on the full diversity of populations.”

Sequencing “the” human genome circa 1990 now seem absurdly simplistic. A genome of 3,054,832 billion DNA base pairs can vary in many ways.

A watermelon genome is a lot smaller than ours – about 350 million DNA base pairs, splayed across 11 chromosomes.

Follow the latest news and policy debates on sustainable agriculture, biomedicine, and other ‘disruptive’ innovations. Subscribe to our newsletter.

[SIGN UP](#)

Better flavor but more susceptible to disease?

The “super-pangenome” for watermelon and its wild relatives, the researchers hope, will uncover beneficial genes lost during domestication. Traits of interest affect seeds; rind thickness; fruit size, shape, texture, and sweetness; and improving disease resistance, which might lessen reliance on agrochemicals.

“We aimed to delve deeper into the genetic variations that make watermelons so diverse and unique. Our findings not only provide insights into the evolutionary journey of watermelons, but also present significant implications for breeding and disease resistance,” said lead author Zhangjun Fei.

The investigators constructed the watermelon super-pangenome using reference (already refined and published) genome sequences as well as resequencing genome data from 547 watermelon accessions for the four species – cultivated watermelon (*Citrullus lanatus*) and its wild relatives *C. mucosospermus*, *C. amarus*, and *C. colocynthis*. In agriculture, an accession is a group of related plant material from a single species collected at one time and place. It attempts to capture the diversity of a particular plant

population.

Analyses of the super-pangenome revealed that many disease-resistance genes in wild species were lost during domestication, as early farmers selected for fruit sweetness, flesh color, and rind thickness. “These beneficial genes could be reintroduced into modern cultivars to breed more resilient watermelon varieties,” noted Fei.

At the root of the sweetness is a tandem duplication of the sugar transporter gene *CITST2*. A tandem duplication is an exact copy of a gene right next to it – like *hippo hippo* in this sentence. Such gene duplication is a driving force in evolution, providing a blank slate that can mutate while the original gene maintains the function. For cultivated watermelon, the extra dose of *CITST2* explains the greater sugar content and sweetness.

The duplicated gene was rare among ancient wild watermelons. But at some point, a mutation in the gene copy imparting extra sweetness must have emerged and astute farmers, noticing the novel trait, adjusted breeding strategies to maximize its transmission. It’s an elegant example of the artificial selection that is the backbone of traditional agriculture building on natural variation and the dynamic nature of DNA.

“The super-pangenome provides a valuable genetic toolkit for breeders and researchers to improve cultivated watermelon,” said Fei. “By understanding the genetic makeup and evolutionary patterns of watermelons, we can develop varieties with enhanced yield, increased disease resistance, and improved adaptability.”

Genetic details, sweetness, and fruit flesh quality

Comparing the genomes revealed the family tree: *C. colocynthis* is the ancestor, its genes jumbled up spawning the other two species. As expected, gene variants with different frequencies among the four species were those that breeding selected.

In addition to the sweet gene *CITST2*, the researchers identified 17 disease-resistance genes that are absent or rare in *C. lanatus* (what we eat), but are present at high frequency in at least one wild relative.

Another forebear to watermelon, perhaps the most recent, may be the [Kordofan melon](#), found in Sudan. It has yellow flesh, a green rind, but is not sweet. It is considered a landrace, less genetically diverse than wild ancestors, yet more diverse than cultivars from plant breeding.

The new work reveals that in modern watermelon’s genome are 123 places, harboring 399 genes, where comparison to Kordofan’s genome indicates signs of domestication. Of these genes, 107 affect fruit quality.

Kordofan, in addition to lacking sweetness, is tough to eat because of a high flesh soluble solids level. That prevents it from melting in the mouth like modern watermelon, which bursts with sweetness from the double dose of the *CITST2* gene. The sweet gene duplication became a fixture in cultivars, as did a high “total soluble solids content” – the fruit is large.

The watermelon ancestor Kordofan also falls short in percent solids, which is mostly sucrose, measured with the Brix standard. For sweet corn the standard is 6 to 24, with 24 being excellent. Watermelons and plums have Brix scores of 9 to 15. Kordofan’s Brix of .2 to 3.2 is far below what’s considered to be a sweet fruit.

Conclude the researchers, “Our *Citrullus* super-pangenome provides insights into watermelon evolution and domestication and serves as a comprehensive resource for researchers and breeders to mine and utilize genes in cultivated and wild watermelon species.”

Coda: Apps

What else has agriculture inadvertently bred out of crops and domesticated animals that we might use? Consider domestic dogs, their genomes selected to the extreme.

Another *DNA Science* post from about a year ago considered borrowing from the genomes of [poodles](#) to fashion the aussiedoodle, bassetoodle, bernedoodle, chipoo, doxiepool, Irish doodle, poochon, rattle, and shihpool, altogether some 50 variations on the theme. The diverse doodles share the trademark tight curly fur, but vary in size, color, head shape, and hard-to-quantify traits such as “confident yet affectionate, but also active and deceptively athletic.” I wonder what has been bred out of their gene pool.

Doodles arise from breeding, of course. But I also wondered how biotechnology might be used in agriculture to more precisely alter the genome to fashion traits of use to us. Consider red meat.

Might we find a biotech way to remove [sialic acid](#) (a type of sugar) from the muscle cell surfaces of beef, pork, and lamb, rendering their flesh less likely to cause inflammation in human carnivores? Sialic acid on our cell surfaces is slightly different than versions on muscle cells of cattle, pigs, and sheep and [others](#) such as goat, deer, and kangaroo. The cells of these animals make an enzyme that dismantles their form of sialic acid, but our cells make a different form that can’t handle the sugar. Instead, the human immune system reacts to sialic acid on these red meats with its inflammatory response. Over time, eating lots of hamburgers, steaks, and ribs raises risks of developing cardiovascular disease, arthritis, and cancer.

Could we engineer-out the molecules of organisms that irk us, such as ragweed, peanuts, wasps and bees, like CRISPR renders infertile malaria-spreading mosquitoes?

Returning to watermelon, I think it’s fascinating that genomics can now be overlaid on the classic plant breeding of agriculture to illuminate the origins of this summer flavor favorite.

Ricki Lewis has a PhD in genetics and is a science writer and author of several human genetics books. She is an adjunct professor for the Alden March Bioethics Institute at Albany Medical College. Follow her at her [website](#) or Twitter [@rickilewis](#)

A version of this article was originally posted at [PLOS](#) and has been reposted here with permission. PLOS can be found on Twitter [@PLOS](#)