

### CROP GENOMICS

## Tomato pan-genome

Nat. Genet. <http://doi.org/c53d> (2019).



Credit: Robert Kneschke / Alamy Stock Photo

The tomato reference and re-sequenced genomes tremendously enhance our understanding of variations that underlie traits and human selection of this important crop. However, existing knowledge based on the reference sequence misses the, also important, nonreference genome. To bridge the gap, Lei Gao, from the Boyce Thompson Institute for Plant Research, Cornell University, and colleagues present a tomato pan-genome assembled from the resequencing data of 725 representative accessions. Their study reveals new nonreference genes and human-selected variations, providing important implications for basic and applied research.

The representative accessions comprise wild species, early domesticates, large-fruited heirlooms, modern elite cultivars and hybrids. Using a ‘map-to-pan’ strategy, the researchers assembled a 1,179 megabase pan-genome that encodes 40,369 predicted protein-coding genes. A total of 4,873 nonreference protein-coding genes were identified, displaying lower expression than reference genes.

The genes in the pan-genome were categorized into four groups — the core group shared by all the accessions, the softcore group shared by more than 99% accessions, shell genes present in 1–99% accessions and cloud genes found in less

than 1% of the accessions. Tomato is exceptionally high in the content of core genes compared to other species.

Analysing the gene presence and absence variations, the researchers found a general trend of gene loss during domestication and improvement. Modern breeding, nevertheless, re-raised gene content in elite inbred lines by intense introgressions of stress resistance alleles from wild species. Genes and promoter sequences under positive or negative selection were also discovered, most of which showed continued selection preference from domestication to improvement.

A rare allele in the promoter of the *TomLoxC* gene was strongly disfavoured during domestication and improvement but its frequency recovered in modern elite lines, likely due to the selection for fruit flavour. This gene regulates apocarotenoid production and thus tomato flavour, as shown by genetic and functional analyses. The heterozygotes of this locus have higher gene expression over both homozygotes in orange-stage fruits, representing a desirable genotype that has been favoured in modern tomato breeding.

---

Jun Lyu

Published online: 29 May 2019  
<https://doi.org/10.1038/s41477-019-0453-5>