

Spotlight

Mobile Transposable Elements Shape Plant Genome Diversity

Saleh Alseekh,^{1,2}
Federico Scossa,^{1,3} and
Alisdair R. Fernie^{1,2,*}

The presence of various types of structural variants, including transposons, make up the major part of the genomic differences among plant species. Two recent papers, Domínguez *et al.* and Alonge *et al.* explore specifically the impact that retrotransposons and other structural variants had on several tomato phenotypes of agricultural importance.

Despite its narrow genetic basis, tomato (*Solanum lycopersicum*) displays wide phenotypic variance in developmental and metabolic traits, with much of this diversity being ascribed to the selection of rare alleles with large effects [1]. However, this viewpoint is challenged by the fact that genome-wide association studies considered so far, include mainly SNPs and short indels (for notable exceptions see [2]), although structural variants (SVs) tend to account for the largest proportion of sequence differences between individuals and cultivars [1]. Two recent papers assess the effect of SVs, and in particular retrotransposons, on tomato gene expression, and explore the consequences of some specific SVs on the manifestation of important agricultural phenotypes [3,4]. Both studies uncover the importance of SVs in understanding the genotype-phenotype equation, and in shaping diversity along tomato domestication and improvement.

The approach taken in both papers was essentially the same, whilst Domínguez *et al.* [3] relied on available resequencing data from over 600 cultivated and wild accessions [5], Alonge *et al.* [4] additionally

generated long read sequencing data for 100 tomato accessions. This new data allowed the authors to establish a PanSV Genome for tomato, with individual accessions having between 1928 and 45 840 SVs with wild species harboring the greatest structural variation [4]. SVs were typically composed of, or generated by, transposons [6], and although their regulatory activities and potential as insertional mutagens have long been known in both animals and plants [7], understanding the way they affect gene regulation and phenotypic diversity on a wider genomic scale has remained controversial.

In order to address this gap, Domínguez *et al.* [3] assessed the set of transposable element (TE) families with recent mobilization activity, by surveying tomato accessions spanning wild, domesticated, and cultivated tomatoes. Analysis of TEs [belonging to class I long terminal repeats (LTR) and non-LTR retroelements, and class II DNA transposons] revealed a total of 6906 nonreference TEs with most being present in a single or few accessions. Analyzing these revealed that the majority of transposon insertion polymorphisms (TIPs) resulted from the mobilization of COPIA-like retrotransposons. Intriguingly, the mobilome fraction varied substantially across tomato domestication. Mobile TE families were found to be more abundant in wild tomato relatives and in the early domesticates of the cerasiforme group, as the result of significant gene flow between these two groups. By contrast, the introduction of founder varieties into European cultivation, resulted in a significant reduction in the number of TE families in modern tomatoes.

Following these analyses, a survey of the TIP landscape revealed that COPIA and many other TE families are found preferentially in or near genes, whilst Gypsy TEs are mainly clustered in the pericentromeric regions. Indeed, experimental evidence indicates that COPIA TEs integrate within environmentally responsive genes [8],

perhaps underlying why they were identified to be conspicuous in the region of stress responsive genes in the *Solanum pennellii* genome [9]. To assess the effect of TIPs on gene expression, RNAseq information from 400 accessions [5] was evaluated, indicating that TIPs within the transcribed protein of genes, had dramatic yet complex effects. Functional analysis of the effects on transcription, revealed potential roles in virus and Phytophthora resistance, as well as shelf life [3]. Similarly, Alonge *et al.* [4] found globally large effects of SVs on gene expression in an RNAseq analysis of cotyledons, roots, and apical meristems, in 23 of the accessions. More important than effects on transcription however, is phenotypic variation. For this purpose, TIP-genome wide association studies (GWAS) were carried out for 17 important agronomic traits, uncovering nine high confidence loci associated with five traits. Most of these TIPs were found to be non-overlapping with other loci associated to SNPs, underlying the importance of screening both SNPs and larger SVs when assessing the genetic basis of quantitative traits (Figure 1). These analyses revealed a new allele of phytoene synthase associated with yellow color, as well as a number of additional TIPs associated with the level of specific metabolites, which were not previously identified by traditional SNP-GWAS [3]. Likewise, the study of Alonge *et al.* [4], identified a P450 gene duplication underlying a fruit weight quantitative trait loci (which was previously thought to be associated with a single SNP) and other SVs, which were required for breeding of the jointless trait, including the insertion of a COPIA TE. A final interesting observation revealed an insertion of a COPIA LTR-retrotransposon, absent in modern cultivars, and associated with the pleasant flowery aroma conferred to heirloom varieties, and associated with high levels of 2-phenylethanol [3,10]. This TE insertion is located within an intron of a gene, putatively encoding an acyl coenzyme (CoA) transferase, involved in the esterification of 2-phenylethanol. Whilst this insertion did not

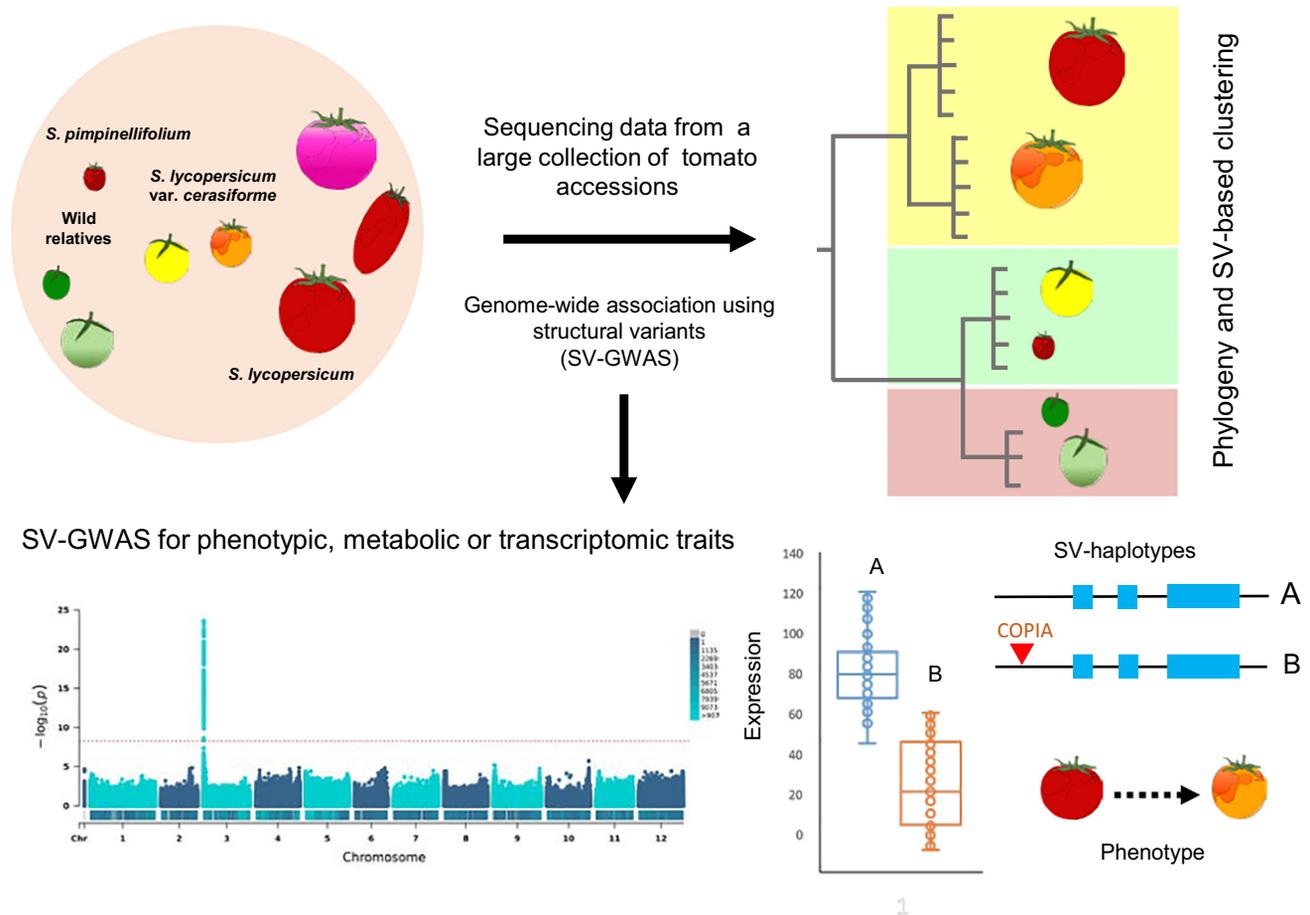


Figure 1. Overview of the Method for Genome-Wide Association Studies (GWAS) for Structural Variants (SVs) in Tomato. DNA-long-read sequencing allows the generation of a catalog of SVs, which is then used to establish the genotype-phenotype relationships during tomato domestication and improvement.

affect the expression levels of the gene, a detailed analysis revealed that it did lead to alternative splicing with truncated transcript isoforms, almost exclusively associated with the intronic COPIA TE insertion. Direct confirmation via gene editing ablation of the TE insertion remains pending. However, the finding that this insertion was absent from wild relatives but present, albeit at low frequency, in *Solanum pimpinellifolium*, suggests that its emergence predated domestication, but was not maintained in modern varieties, which are widely held to have poor flavor.

The two studies by Dominguez *et al.* and Alonge *et al.* represent major advancements in our current knowledge about how SVs, and in particular TIPs, act on rewiring gene

expression, and variation in quantitative traits. Evidence now exists that polymorphisms generated by TEs have been, in some cases, integrated into the regulatory machinery of the host promoting novel cellular phenotypes. At the same time, these SVs have emerged as causal alleles (of large effect size), replacing other minute polymorphisms which were previously, and erroneously, considered to be responsible for the observed phenotypes. With the impressive amount of data reported by these two studies, an attempt can be made to subsequently address other outstanding questions of TE biology. For example, it will be interesting to investigate the historical genesis of TE dispersal across genomes, and discriminate the neutral TEs (or, more

generally, the neutral SVs) from those which showed an immediate adaptive advantage upon integration, or which were later exapted to serve a particular function. These evolutionary studies will help us understand when ancient TE integration events were co-opted to drive the specific gene regulatory mechanisms observed in modern species. Solid evidence now exists from multiple plant species for a preferred association of TE-induced polymorphisms, with genes involved in specific cellular functions (e.g., in stress and environmental responses). Therefore, it will be interesting to evaluate whether this association bias toward specific processes, emerged as a consequence of facilitated integration of TEs in accessible chromatin (during germline

development), and if so, and to which extent, the dispersal of TEs and other structural variants constitutes the result of selective constraints during evolution [11].

¹Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany

²Center of Plant Systems Biology and Biotechnology, Ruski Blvd. 139, Plovdiv, 4000, Bulgaria

³Council for Agricultural Research and Economics (CREA), Research Centre for Genomics and Bioinformatics (CREA-GB), Rome, Italy

*Correspondence:

fernie@mpimp-golm.mpg.de (A.R. Fernie).

<https://doi.org/10.1016/j.tplants.2020.08.003>

© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

References

- Gao, L. *et al.* (2019) The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat. Genet.* 51, 1044–1051
- Akakpo, R. *et al.* (2020) The impact of transposable elements on the structure, evolution and function of the rice genome. *New Phytol.* 226, 44–49
- Dominguez, M. *et al.* (2020) The impact of transposable elements on tomato diversity. *Nat. Commun.* 11, 4058
- Alonge, M. *et al.* (2020) Major impacts of widespread structural variation on gene expression and crop improvement in tomato. *Cell* 182, 145–161.e23
- Zhu, G. *et al.* (2018) Rewiring of the fruit metabolome in tomato breeding. *Cell* 172, 249–261.e12
- Lisch, D. (2013) How important are transposons for plant evolution? *Nat. Rev. Genet.* 14, 49–61
- Chuong, E.B. *et al.* (2017) Regulatory activities of transposable elements: from conflicts to benefits. *Nat. Rev. Genet.* 18, 71–86
- Quadrana, L. *et al.* (2019) Transposition favors the generation of large effect mutations that may facilitate rapid adaptation. *Nat. Commun.* 10, 3421
- Bolger, A. *et al.* (2014) The genome of the stress-tolerant wild tomato species *Solanum pennellii*. *Nat. Genet.* 46, 1034–1038
- Tieman, D. *et al.* (2017) A chemical genetic roadmap to improved tomato flavor. *Science* 355, 391–394
- Schrader, L. and Schmitz, J. (2019) The impact of transposable elements in adaptive evolution. *Mol. Ecol.* 28, 1537–1549