news & views

ADAPTIVE EVOLUTION

Large chromosomal variants drive adaptation in sunflowers

Analysis of genomic and phenotypic data in *Helianthus* species demonstrates that large chromosomal regions contribute to multiple traits associated with distinct ecotypes.

Yongfeng Zhou and Brandon S. Gaut

How does a population adapt to its local environment? If the population is geographically isolated from other populations, adaptation can proceed by well-characterized mechanisms. But what happens when the population exchanges genes with nearby populations that are adapted to different conditions? Alleles from these other populations can swamp locally adapted alleles, slowing or perhaps even halting the adaptive process. A potential evolutionary solution to limit swamping is the formation of local haplotypes that combine adaptive alleles from different loci. If recombination is suppressed among these loci, migrant alleles cannot break up the adaptive combination, giving the haplotype a local fitness advantage[1]. By this process, locally adapted populations are expected to differ from other populations by distinct haplotypes that may span large chromosomal regions. This expectation has been confirmed in a recent study of ecotypic differentiation among sunflower (*Helianthus*) species[2].

Sunflowers are intriguing because they exhibit extensive ecotypic variation across distinct, closely related species. For example, the silverleaf sunflower (*Helianthus argophyllus*) contains two distinct ecotypes in southern Texas, one coastal and the other inland (Fig. 1). The latter must survive and reproduce through elevated summer temperatures, which may explain why it flowers more than two months later than the coastal ecotype. Another dramatic example is the prairie sunflower (*Helianthus petiolaris*), which tends to grow in sandy soils but has ecotypes that grow (quite literally) in sand. *H. petiolaris* has adapted to growth on sand dunes separately in at least two locations: Texas and Colorado. Finally, the common sunflower (*Helianthus annuus*) grows in mesic regions; some populations have extended their range to include deserts and still others are adapted to the rigors of cultivation (Fig. 1).

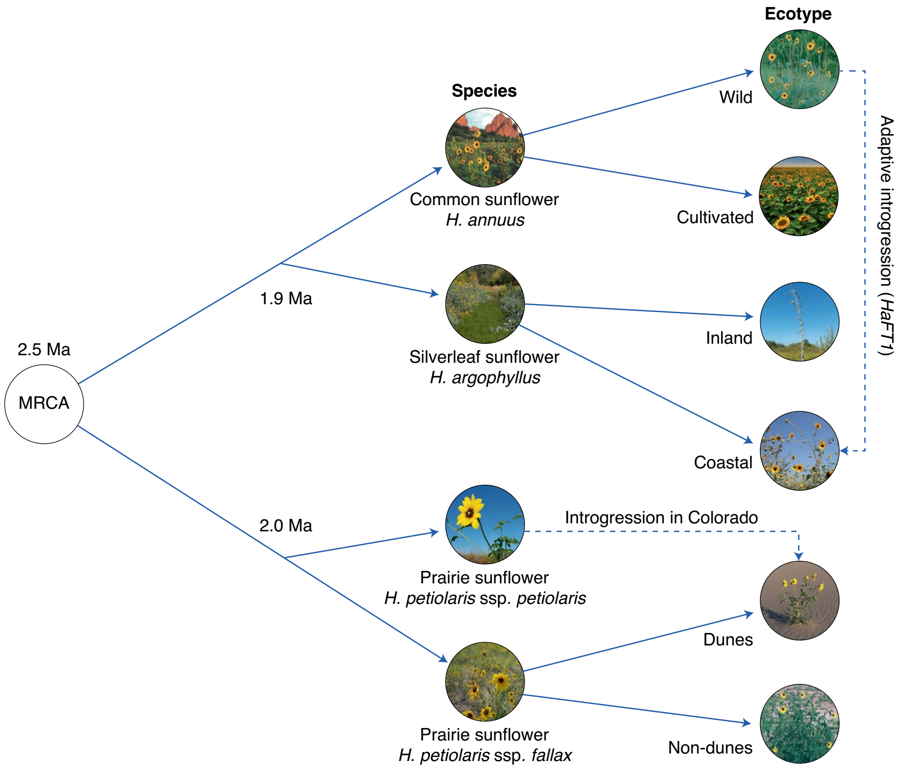

Fig. 1 | A phylogeny of the taxa studied in Todesco et al.[2] that provides the estimated divergence time between species and subspecies on each node. The distinct ecotypes that were studied are shown on the far right. The dotted arrows indicate inferred introgression events based on the analysis of chromosomal blocks. MRCA, most recent common ancestor. Ma, million years ago. Images courtesy of Dylan O. Burge, Nolan C. Kane, Brook T. Moyers, Mariana A. Pascual-Robles and Jason Rick.

To better understand the genetic basis of this extensive ecotypic differentiation, Todesco et al.[2] sampled 151 sunflower populations that included distinct ecotypes among the three species. They grew individuals from each population in a common garden and tracked up to 86 phenotypes per species, including features like flowering time and seed size. After resequencing>1,500 *Helianthus* individuals using a reduced-representation method to remove repetitive genomic fractions, they associated genotypes with phenotypes and environmental variables. The surprising outcome was that they found associations that spanned large genomic blocks; for example, the difference in flowering time between coastal and inland *H. argophyllus* associated with SNPs within a ~10 Mb block on chromosome 6. This block had two major haplotypes, corresponding to early- and late-flowering phenotypes, and contained five of the six sunflower homologues to the *Arabidopsis FLOWERING LOCUS T* (*FT*) locus. The late-flowering haplotype had a deletion of one of those homologues (*H. annuus FT1*), which restored early flowering in an otherwise late-flowering *Arabidopsis thaliana* mutant.

Block-based associations were not unique to contrasts between *H. argophyllus* ecotypes, because three>1 Mb blocks distinguished the dune and non-dune *H. petiolaris* ecotypes, especially their difference in seed size. These findings prompted a clever idea to reverse the order of discovery. Instead of using genotype–phenotype associations to define blocks, the researchers first defined blocks from resequencing data, using a local principal component analysis (PCA) approach that scans windows across the genome[3], and then complemented PCAs with genetic maps and high-throughput chromosome conformation capture data. They identified 37 blocks of>1 Mb that encompassed 4–16% of the genome, depending on the species, and treated each block as a single locus. Most of these blocks were associated with phenotypic or environmental differences between ecotypes. Remarkably,>75% of the associations included more than one trait, suggesting that each block contained locally adapted alleles from multiple loci.

What are these blocks, and how do they arise? A handful of the sunflower blocks were determined to be chromosomal inversions, a particularly well-studied type of structural variant (SV). Inversions suppress recombination, and they have a demonstrable fitness advantage so long as they encompass at least two locally adapted mutations, even when there is gene flow between populations[1]. Previous studies have identified inversions associated with multiple ecologically relevant traits, such as the inversions in wild maize (*Zea mays*)[4] that associate with several climatic and soil factors[5] as well as in yellow monkeyflower (*Mimulus guttatus*) that contribute to differences in life-history characteristics between ecotypes[6]. Although previous studies have shown that large inversions contribute to multiple adaptive traits, a major contribution of the sunflower study is the apparent prevalence of such regions.

We can think of two other explanations for these blocks. One is that they represent emerging 'speciation islands', a phenomenon that combines locally adapted loci in a recombination-suppressed region. In theory, these islands can arise in the face of gene flow with strong selection, and they do not require physical inversion[7]. A related possibility is the presence or absence of lengthy SVs between ecotypes. For such variants, hybrids between ecotypes will be hemizygous over the length of the SV, which should effectively suppress recombination due to a lack of physical pairing. At least one study has shown that substantial regions of plant genomes can be hemizygous[8], suggesting that SVs of this type may be common. Unfortunately, it is currently uncertain how many of the sunflower blocks represent bona fide inversions as opposed to other possibilities, given the limitations of the reduced representation resequencing data.

The sunflower analyses were biased toward detecting large, highly diverged (and therefore old) haplotypes, and these haplotypes proved to have fascinating histories. For example, sequence analyses of the 10 Mb block on chromosome 6 suggested that the early-flowering haplotype was introgressed into *H. argophyllus* from *H. annuus*, perhaps as long as ~2.0 million years ago (Ma) (Fig. 1). Similarly, two haplotype blocks on chromosomes 9 and 11 differed between the dune and non-dune ecotypes of *H. petiolaris* in both Texas and Colorado, suggesting that haplotypes drove convergent adaptation to the dune ecotype (Fig. 1).

An important contribution of Todesco et al.[2] is that it highlights our cavernous gap of knowledge about the prevalence, extent and effect of SVs, and also the challenges to study them effectively. We do know that SVs commonly affect plant phenotypes, because SVs are the causative mutation for at least one-third of known domestication genes[9]. We also know that plant genome sizes can shift dramatically in only a few generations[10], suggesting that large regions of plant genomes may be malleable due to SVs. We are still learning, however, about the prevalence of SVs in populations and their dynamics in local adaptation, speciation and domestication[8]. Population and quantitative genetics analyses promise to provide further insights, but such analyses may require new methods[11] and full resequencing data, preferably with long-read approaches from large samples of individuals. ❐

Yongfeng Zhou ✉ and
Brandon S. Gaut ✉
Department of Ecology and Evolutionary Biology, University of California, Irvine, Irvine, CA, USA.
✉*e-mail: yongfez1@uci.edu; bgaut@uci.edu*

https://doi.org/10.1038/s41477-020-0705-4

References

1. Kirkpatrick, M. & Barton, N. *Genetics* **173**, 419–434 (2006).
2. Todesco, M., Owens, G. L., Bercovich, N. & Légaré, J. S. *Nature* https://doi.org/10.1038/s41586-020-2467-6 (2020).
3. Li, H. & Ralph, P. *Genetics* **211**, 289–304 (2019).
4. Fang, Z. et al. *Genetics* **191**, 883–894 (2012).
5. Aguirre-Liguori, J. A. et al. *Mol. Ecol.* **28**, 2814–2830 (2019).
6. Lowry, D. B. & Willis, J. H. *PLoS Biol.* **8**, e1000500 (2010).
7. Wu, C.-I. *J. Evol. Biol.* **14**, 851–865 (2001).
8. Zhou, Y. et al. *Nat. Plants* **5**, 965–979 (2019).
9. Gaut, B. S., Seymour, D. K., Liu, Q. & Zhou, Y. *Nat. Plants* **4**, 512–520 (2018).
10. Roessler, K. et al. *Nat Plants* **5**, 980–990 (2019).
11. Voichek, Y. & Weigel, D. *Nat. Genet.* **52**, 534–540 (2020).

Competing interests

The authors declare no competing interests.