Figures and figure supplements

Drought adaptation in *Arabidopsis thaliana* by extensive genetic loss-of-function

*J Grey Monroe et al*
Figure 1. Seasonal drought timing varies across the Arabidopsis species range. (A) Examples of home environments for two well-studied Arabidopsis ecotypes (Mojica et al., 2016) from Italy and Sweden, left and right plots respectively, showing historical drought conditions detected using the VHI and drought frequency (VHI <40, NOAA drought classification) by week (line) and season (bars). Arrows mark locally observed flowering dates (Mojica et al., 2016) and gray bars highlight the typical reproductive growing season used to quantify a drought-timing index. (C) Variation in historical drought timing experienced at the home environments of Arabidopsis ecotypes across the species range (figure supplement). Large values indicate environments where spring droughts occur more frequently than summer drought (i.e. where the frequency of drought decreases over the course of the typical reproductive growing season) and vice versa.

DOI: https://doi.org/10.7554/eLife.41038.003
Figure 1—figure supplement 1. Arabidopsis ecotypes are distributed across satellite-detected drought timing gradients. Historical patterns in drought conditions were calculated from the Vegetative Health Index (VHI, Figure 1A) and converted into a drought-timing index (Figure 1B and C). Large values of this index indicate environments where spring droughts occur more frequently than summer drought (i.e. where the frequency of drought decreases over the course of the reproductive growing season) and vice versa (seasonal drought frequency map data available at greymonroe.github.io/data).

DOI: https://doi.org/10.7554/eLife.41038.004
LoF alleles share associations between drought timing and flowering time, exhibit evidence of positive selection. (A) Visualization of the frequency of LoF alleles across environments in genes associated to summer drought timing environments (7 genes).

Figure 2 continued on next page

Monroe et al. eLife 2018;7:e41038. DOI: https://doi.org/10.7554/eLife.41038
Figure 2 continued

(upper) or spring drought environments (lower). Darker lines indicate the mean across genes. (B) Contrasting flowering times between ecotypes with functional versus LoF alleles in genes associated with earlier (upper) or later (lower) flowering time phenotypes. (C) Overlap and relationships between the strength of LoF allele associations in genes associated with summer drought and earlier flowering, and (D) spring drought and later flowering. (E) Increased frequencies of independent LoF alleles in genes associated with drought timing and/or flowering time compared to genes without detected associations (t-test, p = 3.4 × 10⁻⁷), a signature of recurrent mutation accompanied by positive selection (Pennings and Hermisson, 2006).

DOI: https://doi.org/10.7554/eLife.41038.005
Figure 2—figure supplement 1. P values of LoF allele associations. Observed vs. expected P values, created using GWASTools in R (Gogarten et al., 2012), for associations between drought timing and (A) LoF alleles observed in Arabidopsis ecotypes and (B) randomized LoF genotypes with the same
Allele frequencies. Observed vs. expected p values for associations between flowering time and (C) LoF alleles observed in Arabidopsis ecotypes and (D) randomized LoF genotypes with the same allele frequencies. Relationship between LoF allele associations with drought timing and flowering time for (E) actual (C = A, $r^2 = 0.48$) and (F) randomized genes (D = B, $r^2 = 0.01$). The P values shown have not yet been corrected for multiple testing and are log$_{10}$ transformed. Red lines in A-D represent $y = x$ line.

DOI: https://doi.org/10.7554/eLife.41038.006
Figure 2—figure supplement 2. Signatures of selection on LoF genes identified differ from null expectations. (A) Contrasts (t-test, $\alpha = 0.05$) between genes identified with LoF alleles associated to drought timing and/or flowering time (colors correspond to Figure 2C and D, boxplots visualized at $\pm 1.5$ times the data interquartile range) and the genomic background (light gray), as well as genes having LoF alleles but without observed associations (dark gray) for the ratio of non-synonymous ($P_N$) and synonymous polymorphisms ($P_S$) among A. thaliana ecotypes and (B) the ratio of non-synonymous ($D_N$) and synonymous divergence ($D_S$) from A. lyrata. (C) Contrasts (t-test, $\alpha = 0.05$) between (log$_{10}$) global frequency of LoF alleles in genes identified with LoF alleles associated to drought timing and/or flowering time and genes with LoF alleles but without observed associations for the global frequency of LoF alleles and (D) the number of (log$_{10}$) unique LoF alleles. The corresponding average frequencies of unique LoF alleles for genes are shown in Figure 2E.

DOI: https://doi.org/10.7554/eLife.41038.007
Figure 2—figure supplement 3. LoF alleles are not broadly overabundant in Arabidopsis ecotypes originating from spring drought environments or flowering later. (A) The frequency of LoF alleles across environments (sliding window plot) in random genes. The darker line indicates the mean across genes. The distribution of LoF alleles in these random genes contrasts with LoF alleles in genes associated to drought timing, which are overwhelming associated to spring drought environments (Figure 2A). (B) Flowering times compared between ecotypes with functional versus LoF alleles in random genes. The phenotypic differences predicted by these random genes contrasts with LoF alleles in those associated to flowering time, which are overwhelming associated to later flowering time (Figure 2B).

DOI: https://doi.org/10.7554/eLife.41038.008
Figure 3. Widespread LoF contributing to later flowering time evolution. (A) Genomic map of 214 candidate genes with associations between LoF alleles and spring drought environments and/or later flowering time phenotypes. (B–E) Examples of the geography and flowering times among Arabidopsis ecotypes of LoF alleles in candidate genes including; (B) a previously unstudied rhamnogalacturonate lyase, (C) a cyclin linked to later flowering in prior knockout experiments (Cui et al., 2007), (D) members of the drought-responsive Nramp2 (Qin et al., 2017) (E) and RmlC-like cupin (Aghdasi et al., 2012) protein families. (F) Later flowering time in ecotypes predicted by the accumulation of LoF alleles across all candidate genes. The line shows the best fitting model. Color scale of points reflects proportion of total LoF in ecotypes that are candidate genes (darker points = greater proportion) (G) Experimental validation of hypothesized later flowering time in T-DNA knockout lines of candidate genes compared to the wild type genotype.

DOI: https://doi.org/10.7554/eLife.41038.009