OXFORD GENETICS

GENETICS, 2023, 224(2), iyad074
https://doi.org/10.1093/genetics/iyad074
Advance Access Publication Date: 24 April 2023
Investigation

Weaker selection on genes with treatment-specific expression consistent with a limit on plasticity evolution in *Arabidopsis thaliana*

Miles Roberts,[1,*] Emily B. Josephs[2,3,*]

[1]Genetics and Genome Sciences Program, Michigan State University, East Lansing, MI 48824, USA
[2]Department of Plant Biology, Michigan State University, East Lansing, MI 48824, USA
[3]Ecology, Evolution, and Behavior Program, Michigan State University, East Lansing, MI 48824, USA

*Corresponding author: Genetics and Genome Sciences Program, Michigan State University, East Lansing, MI 48824, USA. Email: milesdroberts@gmail.com; *Corresponding author: Department of Plant Biology, Michigan State University, East Lansing, MI 48824, USA. Email: josep993@msu.edu

Abstract

Differential gene expression between environments often underlies phenotypic plasticity. However, environment-specific expression patterns are hypothesized to relax selection on genes, and thus limit plasticity evolution. We collated over 27 terabases of RNA-sequencing data on *Arabidopsis thaliana* from over 300 peer-reviewed studies and 200 treatment conditions to investigate this hypothesis. Consistent with relaxed selection, genes with more treatment-specific expression have higher levels of nucleotide diversity and divergence at nonsynonymous sites but lack stronger signals of positive selection. This result persisted even after controlling for expression level, gene length, GC content, the tissue specificity of expression, and technical variation between studies. Overall, our investigation supports the existence of a hypothesized trade-off between the environment specificity of a gene's expression and the strength of selection on said gene in *A. thaliana*. Future studies should leverage multiple genome-scale datasets to tease apart the contributions of many variables in limiting plasticity evolution.

Keywords: plasticity, partial correlations, evolutionary rates, environment-specific expression

Introduction

Organisms must cope with ever-changing environmental conditions to survive and reproduce. If these changes in condition cannot be avoided or escaped, phenotypes that respond to environmental variation through phenotypic plasticity may be adaptive. For example, under low light, the same *Arabidopsis thaliana* genotype will produce more or larger leaves to capture more energy for photosynthesis (Pigliucci and Kolodynska 2002). Plastic responses are partly controlled through differential gene expression between environments (Scheiner 1993; Schlichting and Smith 2002). Understanding the evolution of these condition-specific expression patterns could help reconcile the diversity of plastic responses observed in nature and engineer organisms to overcome environmental challenges.

However, not all organisms can respond plastically to environmental change, so it is crucial to understand the processes that constrain plasticity (Van Kleunen and Fischer 2005). These constraints are usually characterized as either costs, where plasticity reduces fitness in some way, or limits to the evolution or maintenance of plasticity (DeWitt *et al.* 1998). Decades of research has attempted to measure the costs associated with plasticity (reviewed in Schneider 2022) but studies often fail to detect costs or find costs that are weak or restricted to certain environments (Van Kleunen and Fischer 2005; Van Buskirk and Steiner 2009; Auld *et al.* 2010). Theory also predicts that there will be strong selection to alleviate costs (Murren *et al.* 2015). Thus, limits may be more important than costs in shaping the evolution of plasticity.

Recent work suggests that relaxed selection can limit plasticity evolution (Snell-Rood *et al.* 2010; Murren *et al.* 2015). For instance, one hypothesis posits that genes are often under selection for environment-specific expression to minimize deleterious pleiotropy (Snell-Rood *et al.* 2010; McGuigan *et al.* 2014; Huber *et al.* 2017). However, narrowing the range of environments where a gene is expressed also reduces the opportunity for negative selection to act on deleterious mutations in the gene (Kawecki 1994; Whitlock 1996; Van Dyken and Wade 2010). The accumulation of deleterious mutations could then cancel out any selective benefits of the environment-specific expression pattern. Thus, a trade-off arises between a gene's degree of environment-specific expression and the strength of negative selection acting on said gene. If we assume that environment-specific expression generally contributes to phenotypic plasticity, then this trade-off would potentially limit the maintenance of plasticity (Kawecki 1994; Snell-Rood *et al.* 2010). Whether such a trade-off exists has not yet been tested, but the deposition of expression data from hundreds of experimental treatments across hundreds of labs into public repositories now enables approximating environment

Received: March 06, 2023. **Accepted:** April 07, 2023

specificity as treatment specificity and linking treatment-specific expression to the rate of evolution.

One challenge in studying the relationship between treatment specificity and protein evolution is that many factors influence evolutionary rates (for review, see Rocha 2006; Gaut *et al.* 2011; Koonin 2011; Zhang and Yang 2015) and these factors are hard to disentangle. A protein's expression level is often considered the best predictor of its evolutionary rate (Rocha 2006)—a result observed across all domains of life (Zhang and Yang 2015) and sometimes considered a "law" of genome evolution (Koonin 2011). Among multicellular organisms, the degree of tissue specificity in expression is also generally predictive of evolutionary rates (Duret and Mouchiroud 2000; Winter *et al.* 2004; Zhang and Li 2004; Larracuente *et al.* 2008; Slotte *et al.* 2011; Bush *et al.* 2015; Mukherjee *et al.* 2016; Groen *et al.* 2020; Huang 2022). Additional factors that also influence evolutionary rates include exon edge conservation (Bush *et al.* 2015), mutational bias (Wang *et al.* 2004; Ossowski *et al.* 2010), gene length (Mukherjee *et al.* 2016), gene age (Moutinho *et al.* 2022), GC content (Zhang *et al.* 2002; Mukherjee *et al.* 2016), expression stochasticity (Groen *et al.* 2020), involvement in general vs specialized metabolism (Mukherjee *et al.* 2016), identity as a regulatory or structural gene (Wheeler *et al.* 2022), recombination rate (Langley *et al.* 2012), codon-bias (Betancourt and Presgraves 2002), mating system (Wright *et al.* 2002; Glémin 2007; Payne and Alvarez-Ponce 2018), gene compactness (Larracuente *et al.* 2008; Mukherjee *et al.* 2016), co-expression or protein–protein interaction network connectivity (Alvarez-Ponce and Fares 2012; Alvarez-Ponce *et al.* 2017; Josephs *et al.* 2017; Mähler *et al.* 2017; Masalia *et al.* 2017), gene body methylation (Takuno and Gaut 2012), metabolic flux (Colombo *et al.* 2014), protein structure (Lin *et al.* 2007), essentiallity (Nembaware *et al.* 2002; Yang *et al.* 2003; Davis and Petrov 2004), and even plant height (Lanfear *et al.* 2013). This overabundance of possible explanatory variables suggests that massive genome-scale datasets and careful statistical analysis are required to tease out the influence of treatment-specific expression on evolutionary rates.

To investigate the influence of treatment-specific expression on evolutionary rates, we compiled a dataset of gene expression data across over 200 treatments from over 300 peer-reviewed studies in *A. thaliana*. We annotated RNA-sequencing runs from these studies using standardized ontologies, then processed all of them with the same pipeline. Finally, we combined the resulting gene expression matrix with estimates of selection based on within-species polymorphism and between-species divergence to investigate whether genes with treatment-specific expression were under weaker negative selection.

Materials and methods

RNA-seq run annotation

We amassed an initial set of RNA-seq runs from the Sustech Arabidopsis RNA-seq database V2 (Zhang *et al.* 2020) (http://ipf.sustech.edu.cn/pub/athrdb/) excluding any samples not associated with a publication or lacking a tissue type label. On 2022 May 24, we also downloaded all run metadata from the Sequence Read Archive (SRA) returned by the following search term: ("Arabidopsis thaliana"[Organism] AND "RNA"[Source]) OR ("Arabidopsis thaliana"[Organism] AND "RNA-Seq"[Strategy]) OR ("Arabidopsis thaliana"[Organism] AND "TRANSCRIPTOMIC" [Source]). All SRA runs were linked to their associated publications, if possible, using Entrez. Any SRA run numbers that we could not link to a PUBMED ID or DOI were omitted. We then manually removed all SRA runs that originated from transgenic, mutant, hybrid, grafted, cell culture, polyploid, or aneuploid samples based on information in the SRA metadata and associated publications. Runs from any naturally-occurring *A. thaliana* accession were included. We also omitted SRA runs that focused on sequencing non-coding RNA (ncRNA-seq, miRNA-seq, lncRNA-seq, sRNA-seq, etc.). After applying these criteria, any bioprojects with 8 or fewer SRA run numbers remaining were also omitted.

All runs were labeled with treatment and tissue-type descriptions using the Plant Experimental Conditions Ontology (PECO) and the Plant Ontology (PO) (Cooper *et al.* 2018), respectively, based on information in their associated publications and SRA metadata. In our analysis, control exposure was defined as long-day conditions (12 h light exposure or longer, but not constant light) and growing temperatures in the range of 18°–26°, inclusive, without explicit application of stress or nutrient limitation. Warm treatments were defined as 27° or higher, while cold treatments were defined as 17° or lower. Any studies that did not report both day length and growing temperature were omitted. Any runs that could not be linked to treatments based on their annotations in the SRA or Sustech databases were also omitted. Treatment with polyethylene glycol (PEG) was categorized as drought exposure. Samples from plants that were recovering from stress were categorized according to the growth conditions of the recovery state instead of the stressed state. When appropriate, we labeled samples with multiple PECO terms. For example, a sample that was subjected to both heat stress and high light stress would get two PECO terms (one for each stress) and be treated separately from samples subjected to only heat stress or only light stress. Tissue-type labels were eventually collapsed to the following categories: whole plant, shoot, root, leaf, seed, and a combined category of flower and fruit tissues. The flower and fruit tissue categories were combined because of their developmental relationship and small size relative to the other categories. In the end, we had a dataset of 24,101 sequencing runs from 306 published studies.

RNA-seq run processing

All RNA-seq runs were processed using the same workflow to remove the effects of bioinformatic processing differences between studies on expression level. First, runs were downloaded using the SRA toolkit (v2.10.7), but 90 runs were not publicly available and thus failed to download. All successfully downloaded runs were trimmed using fastp v0.23.1 (Chen *et al.* 2018), requiring a minimum quality score of 20 and a minimum read length of at least 25 bp (-q 20 -l 25). Trimming results were compiled using multiqc v1.7 (Ewels *et al.* 2016). All trimmed runs were then aligned to a decoy-aware transcriptome index made by combining the primary transcripts of the Araport11 genome annotation (Cheng *et al.* 2017) with the *A. thaliana* genome in salmon v1.2.1 (Patro *et al.* 2017) using an index size of 25 bp. The salmon outputs of each run were then combined with a custom R script to create an gene-by-run expression matrix. We omitted 423 runs with a mapping rate <1%, 215 runs with zero mapped transcripts, and 18 genes with zero mapped transcripts across all runs from further analysis. We note that although this cut-off does not exclude samples with more modest mapping rates (e.g. 20–60%) the choice to include these samples was to avoid removing large chunks of data as "outliers" and analyzing only those samples that conform to our expectations.

Whole-genome sequence data processing

We downloaded whole-genome sequencing data for 1,135 *A. thaliana* accessions from the 1,001 genomes project panel (SRA project

SRP056687) (Alonso-Blanco *et al.* 2016) using the SRA toolkit. All runs were trimmed using fastp (Chen *et al.* 2018), requiring a minimum quality score of 20 and a read length of at least 30 bp (-q 20 -l 30). Trimmed reads were then aligned to the *A. thaliana* reference genome using BWA v0.7.17 (Li and Durbin 2009). The alignments were sorted and converted to BAM format with SAMTOOLS v1.11 (Danecek *et al.* 2021), then optical duplicates were marked with picardtools v2.22.1. Haplotypes were called for each accession, then combined and jointly genotyped with GATK v4.1.4.1 assuming a sample ploidy of 2, heterozygosity of 0.001, indel-heterozyogsity of 0.001, and minimum base quality score of 20. Invariant sites were included in the genotype calls with the–include-non-variant-sites option. All calls were restricted to only coding sequence (CDS) regions based on the Araport11 annotation by supplying a BED file of CDS coordinates made with bedtools (v2.29.2). Following Korunes and Samuk (2021), variant and invariant sites were filtered separately using both GATK and vcftools v0.1.15 (Danecek *et al.* 2011). Variant sites were filtered if they met any of the following criteria: $QD < 2$, $QUAL < 30$, $MQ < 40$, $FS > 60$, $HaplotypeScore > 13$, $MQRankSum < -12.5$, $ReadPosRankSum < -8.0$, mean depth < 10, mean depth > 75, missing genotype calls > 20%, being an indel, or having more than two alleles. In the end, 1,915,859 variant sites across all coding sequences were retained for further analysis. Invariant sites were filtered if they met any of the following criteria: QUAL > 100, mean depth < 10, mean depth > 75, missing genotype calls > 20%. Finally, variant sites were annotated using snpEff (Java v15.0.2) (Cingolani *et al.* 2012b) and variants labeled as either missense or synonymous were separated into different files using SnpSift (Cingolani *et al.* 2012a).

Selection estimated from between-species divergence

We identified 1:1 orthologs between the primary transcripts of *A. thaliana* and *Arabidopsis lyrata* with Orthofinder v2.5.4 (Emms and Kelly 2019). For each 1:1 ortholog, we aligned their protein sequences with MAFFT L-INS-I v7.475 (Katoh and Standley 2013), then converted the protein alignments to gapless codon-based alignments using pal2nal v14 (Suyama *et al.* 2006). Using the gapless codon-based alignments, we estimated dN/dS using the Nei and Gojobori (1986) method implemented as a custom Biopython v1.79 script and implemented through the codeml program in the PAML package v4.9 (Yang 2007). Unlike codeml, the custom Biopython script also returns counts of nonsynonymous (N) and synonymous sites (S) within each gene as described in Nei and Gojobori (1986), which we later used to calculate nucleotide diversity per nonsynonymous site (π_N) and per synonymous site (π_S). Before proceeding with more analyses, we confirmed that our estimates of dN and dS were consistent between our Biopython script and codeml (Supplementary Fig. S5, Pearson correlations $dN: \rho = 0.9998$, $dS: \rho = 0.9809$). The outputs of the Biopython script were used in all subsequent analyses.

Selection estimated from within-species polymorphism

Nucleotide diversity at nonsynonymous sites

Nucleotide diversity (π) was calculated for each gene with pixy v1.2.3.beta1 (Korunes and Samuk 2021) three times: once using all sites (both variant and invariant), once using missense sites plus invariant sites, and once using synonymous sites plus invariant sites. These estimates were then converted to π, π_N, and π_S, respectively, by first multiplying the per site estimate output from pixy by the number of sites included in the analysis. Then, to get π_N and π_S, the values from analyses of missense plus invariant, and synonymous plus invariant sites were divided by the N and S values for each gene, respectively, as determined by the method in Nei and Gojobori (1986).

Tajima's D

We next calculated Tajima's D for each gene. First, we calculated π and Watterson's Theta (θ_W) for each variant site i within a gene (π_i and θ_{Wi}, respectively). In this case, π_i was calculated as:

$$\pi_i = \left(\frac{n_i}{n_i - 1}\right)\left(1 - \sum_{j=1}^{2} p_{ij}^2\right) \tag{1}$$

Where n_i is the number of sequenced chromosomes with non-missing genotypes for variant i, p_{i1} is the frequency of the reference allele, and p_{i2} is the frequency of the alternative allele. Then, θ_{Wi} was calculated as:

$$\theta_{Wi} = \frac{1}{a_i} \tag{2}$$

Where a_i is:

$$a_i = \sum_{j=1}^{n_i - 1} \frac{1}{j} \tag{3}$$

This calculation of θ_{Wi} is equivalent to the usual calculation of θ_W with the number of segregating sites set to one. Next, the variance in Tajima's D was calculated for each site as:

$$\mathrm{Var}(\pi_i - \theta_{Wi}) = \frac{\frac{n_i + 1}{3(n_i - 1)} - \frac{1}{a_i}}{a_i} \tag{4}$$

This is equivalent to equation 38 in Tajima (1989) with the number of segregating sites set to one.

Finally, the results of the above calculations were combined in the following formula:

$$D_i = \frac{\pi_i - \theta_{Wi}}{\sqrt{\mathrm{Var}(\pi_i - \theta_{Wi})}} \tag{5}$$

To get Tajima's D for each gene, we then averaged across the D_i values for all the variant sites within a gene.

Direction of selection (DoS)

Counts of nonsynonymous and synonymous polymorphisms within each gene (P_N and P_S, respectively) were determined with bedtools (v2.29.2). The number of nonsynonymous and synonymous differences (D_N and D_S, respectively) between *A. thaliana* genes and their 1:1 *A. lyrata* orthologs, if present, were estimated during the process of calculating dN/dS in Biopython as described above. These values were then used to calculate the direction of selection (DoS) (Stoletzki and Eyre-Walker 2011) as follows:

$$\mathrm{DoS} = \frac{D_N}{D_N + D_S} - \frac{P_N}{P_N + P_S} \tag{6}$$

We chose this metric, as opposed to the proportion of amino acid substitutions driven by positive selection (α), because it is less biased than α (Stoletzki and Eyre-Walker 2011) and was successfully used in studies similar to ours (Paape *et al.* 2013). Furthermore, we found that α often returns uninterpretable negative values when applied to *A. thaliana*, perhaps because of an excess of slightly deleterious polymorphisms (Nordborg *et al.* 2005) due to their predominantly selfing mating system (Charlesworth 1994).

Treatment specificity

Treatment specificity (τ) was estimated separately for runs from each tissue type using the following formula (Yanai *et al.* 2005):

$$\tau = \frac{\sum_{i=1}^{N} 1 - \frac{x_i}{\max x}}{N - 1} \tag{7}$$

Where *x* is the vector of average expression values of a gene in each treatment category, measured in transcripts per million (TPM), and where *N* is the number of treatment categories. Dividing by *N*–1 means that τ varies between zero and one, where zero indicates no specificity and one indicates exclusive specificity to a single treatment. We used this metric of specificity because it is consistently more robust than others (Kryuchkova-Mostacci and Robinson-Rechavi 2017) and is normalized by the number of treatments included, making it comparable across datasets. We also applied the same formula to calculate tissue specificity in several different treatment conditions.

Simulating correlations between average expression and specificity index

Average expression level and measures of expression specificity are correlated by definition because genes with more treatment/tissue-specific expression will have lower average expression across all treatment/tissue categories. We ran two simulations to better illustrate the factors driving the correlation between average expression and the specificity index, τ. In both simulations, we generated 1,000 random matrices, where each element x_{ij} represented the expression of gene *i* in experiment *j*, by sampling from a zero-inflated negative binomial distribution:

$$x_{ij} \sim ZINegBinom(N, p_1, p_2) \tag{8}$$

Where the size and probability parameters of the negative binomial component were $N = 100$ and $p_1 = 0.1$, respectively, while the probability of an expression value being non-zero was $p_2 = 0.4$. All matrices included five groups of columns, with five columns per group, representing replicates of tissue/treatment groups. For both simulations, we averaged across columns within each group to simulate the calculation of tissue/treatment-wide averages. We then applied the formula for τ across the rows of this averaged matrix to get expression specificity. In one simulation, we calculated expression level by averaging across the rows of the expression matrix. In a second simulation, we excluded experiments where a gene was not expressed ($x_{ij} = 0$) from the calculation of average expression.

Average expression, length, GC content, family size

Calculating the average expression of each gene was a three-step process. First, we averaged together runs with matching SRA experiment IDs because these runs represented technical replicates of the same biological sample and treatment conditions. Second, we partitioned our gene-by-experiment expression matrix by the tissue type each sample came from. Finally, for each tissue type's expression matrix, we averaged across all of the expression values of each gene across all experiments, excluding values <5 transcripts per million (TPM). We excluded values <5 TPM from the average expression calculation to avoid a high correlation between average expression and treatment specificity, as has been reported in previous studies (Slotte *et al.* 2011). This high correlation occurs because an environment-specific gene will by definition also have low average expression across environments it is rarely expressed in. Furthermore, we excluded values <5 TPM to avoid including small expression values that could be artifacts of alignment error.

The length and GC content of each gene was measured using the bedtools nuc command (v2.29.2) and included each gene's introns and untranslated regions when present. We included introns and untranslated regions in the estimate of gene length because they play important roles in determining rates of protein evolution (Castillo-Davis *et al.* 2002; Eisenberg and Levanon 2003). Finally, the family size for each gene was estimated as the number of *A. thaliana* genes in their respective orthogroups output by OrthoFinder.

Partial correlation analysis

Not all treatment-tissue combinations were sampled in the overall RNA-seq dataset, causing confounding between the treatment and tissue labels. We resolved this in two ways. First, we subset the data to only the treatment conditions where all tissue types were represented. Second, we subset the data by tissue type and analyzed each subset separately. For each subset, we calculated partial spearman correlations between treatment specificity and our measures of selection (*dN*, π_N, Tajima's D, and *DoS*) after accounting for average expression (excluding values TPM < 5), gene length, and GC content using the ppcor R package (Kim 2015). For partial correlation analyses involving π_N and Tajima's D, we also controlled for gene family size. We did not account for gene family size in partial correlation analyses involving *dN* or *DoS* because these metrics apply to only genes with one family member in this study. When calculating partial correlations involving *dN*, we excluded any genes with saturating divergence ($dS > 1$). All statistical analyses and data visualizations used R v4.0.3 and used color palettes in the scico R package (Crameri 2018; Pedersen and Crameri 2022).

Surrogate variable analysis

We recalculated treatment specificity and repeated all partial correlation analyses after correcting each data subset for technical between-experiment variation (i.e. batch effects), following an approach from (Fukushima and Pollock 2020). Batch effects include variables that influence gene expression measurements but are not of interest to this study, such as the sequencing platform and the library prep protocol used in each experiment. First, with our data already subset by tissue type, we further subset to only include treatments with RNA-seq runs from at least two studies. This minimizes confounding between-treatment variation with the technical between-experiment variation we aimed to account for. We then applied surrogate variable analysis (SVA) using the svaseq() function within the SVA package (Leek and Storey 2007) to each of these subsets. Briefly, SVA models gene expression as:

$$x_{ij} = \mu_i + f(y_i) + e_{ij} \tag{9}$$

Where x_{ij} is the expression of gene i in experiment j, μ_i is the average expression of gene i across all experiments, and y_i is the value of a predictor variable of interest for gene i. Furthermore, $f(y_i)$ gives the deviation of gene i from its average expression based on the value of y_i and e_{ij} is the residual error. SVA takes this model and partitions the residual variance, e_{ij}, into:

$$x_{ij} = \mu_i + f(y_i) + \sum_{\ell=1}^{L} \gamma_{\ell i} g_{\ell j} + e^*_{ij} \tag{10}$$

Where $\sum_{\ell=1}^{L} \gamma_{\ell i} g_{\ell j}$ gives the summed effects of L unmodeled variables ($g_{\ell j}$) for each gene and e^*_{ij} gives the gene-specific noise in expression. SVA does not attempt to estimate what the unmodeled variables influencing expression are, but rather find a set of vectors (the surrogate variables) that span the same space as $\boldsymbol{g}$:

$$x_{ij} = \mu_i + f(y_i) + \sum_{k=1}^{K} \lambda_{ki} h_{kj} + e^*_{ij} \tag{11}$$

Where each $\boldsymbol{h_k}$ is a surrogate variable and each $\boldsymbol{\lambda_k}$ gives the effects of each surrogate variable on gene expression. For our analyses, our predictor variable y_i was treatment type. To get a measure of expression where the effects of surrogate variables are removed, we then subtracted off the effects of surrogate variables from both sides of the above equation.

$$x_{ij} - \sum_{k=1}^{K} \lambda_{ki} h_{kj} = \mu_i + f(y_i) + e^*_{ij} \tag{12}$$

Where $x_{ij} - \sum_{k=1}^{K} \lambda_{ki} h_{kj}$ gives us our expression values accounting for the effects of surrogate variables. The net result here is a reduction in the amount of unexplained or seemingly stochastic variation in expression because sources of variation have been attributed to "surrogates" that span the same space as real batch variables. We also conducted principal component analysis in R before and after SVA to verify the removal of batch effects.

Results

Summary of tissue differentiation, treatment specificity, and selection in overall dataset

To understand how treatment specificity of gene expression affects evolutionary rates of proteins, we queried the Sequence Read Archive for all *A. thaliana* RNA-seq experiments published before May 2022. We then annotated these experiments with standardized tissue and treatment ontology terms, manually filtered the dataset, and then processed all RNA-seq runs with a standardized pipeline. The number of sequencing experiments associated with each combination of tissue and treatment labels is summarized in Supplementary Table S1. Overall, the most sampled tissue category was leaf (4,642 experiments) followed by root (3,348 experiments), whole plant (2,492 experiments), seed (1,866 experiments), shoot (1,106 experiments), then fruit and flower (266 experiments). The four most sampled treatment categories were control (5,701 experiments), cold air exposure (675 experiments), short day length (561 experiments), and short day length plus *Botrytis cinerea* exposure (407 experiments). Any sequencing runs that shared an SRA experiment ID were averaged to produce individual gene expression values for each SRA experiment.

We first looked at the distribution of mapping rates across all RNA-seq runs. The median mapping rate was 72.39% (Supplementary Fig. S1) and we excluded runs with a mapping rate <1% from further analyses. We next performed a principal components analysis (PCA) on the expression matrix and observed strong differentiation between root and non-root tissues along PC2 (Fig. 1). We also observed that nearly all genes had some degree of treatment specificity in their expression (Fig. 2a, Supplementary Fig. S3). Furthermore, only a small proportion of genes had strong signatures of selection based on dN/dS, π_N/π_S, DoS, or Tajima's D (Fig. 2, b to d, Supplementary Fig. S2). The treatment specificity of expression was lower on average in flower and fruit tissue compared with the other tissues (Supplementary Fig. S3). However, tissue specificity did not vary widely depending on the treatment condition (Supplementary Fig. S4).

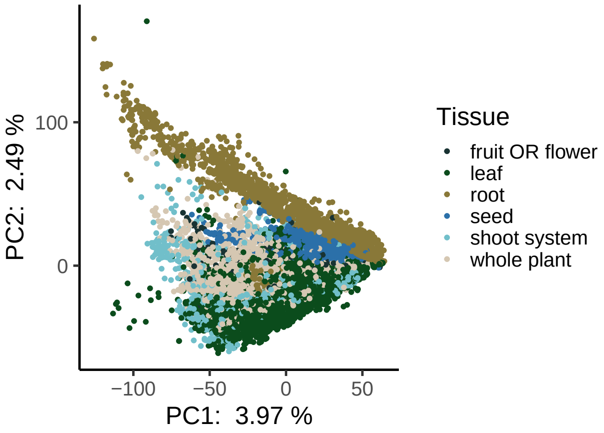

Fig. 1. Principal components analysis of all expression data. Each point represents a different RNA-seq experiment and is colored by its associated tissue type. Experiments from all treatment conditions are included in this analysis. The percent values on the axes represent the percent variation explained by each principal component. Plotting order was randomized to avoid overplotting.

Omitting samples with low expression disentangles expression level and specificity

Genes that are only expressed in one treatment or tissue will, by definition, have low mean expression across all environments or tissues (Wright *et al.* 2004). Thus, we sought a method of calculating expression level that was independent of treatment specificity. To better understand the relationship between average expression and treatment specificity, we calculated correlations between treatment specificity and expression level while either including or excluding low expression values (TPM <5) on our real RNA-seq dataset. We found that excluding low expression values decreased the correlation between average expression and treatment specificity in leaf tissue samples (Fig. 3) and other tissues (Supplementary Figs. S34–S38) and replicated the result by simulating gene expression matrices (Supplementary Fig. S39). Thus, for all later partial correlation analyses (see next section) we quantified each gene's average expression after dropping experiments where the gene was not expressed (TPM < 5).

Treatment specificity correlates with levels of nonsynonymous diversity and divergence in genes

We next calculated partial correlations between treatment specificity and measures of selection after controlling for average expression, gene length, GC content, and tissue specificity in expression. These partial correlations were calculated separately for

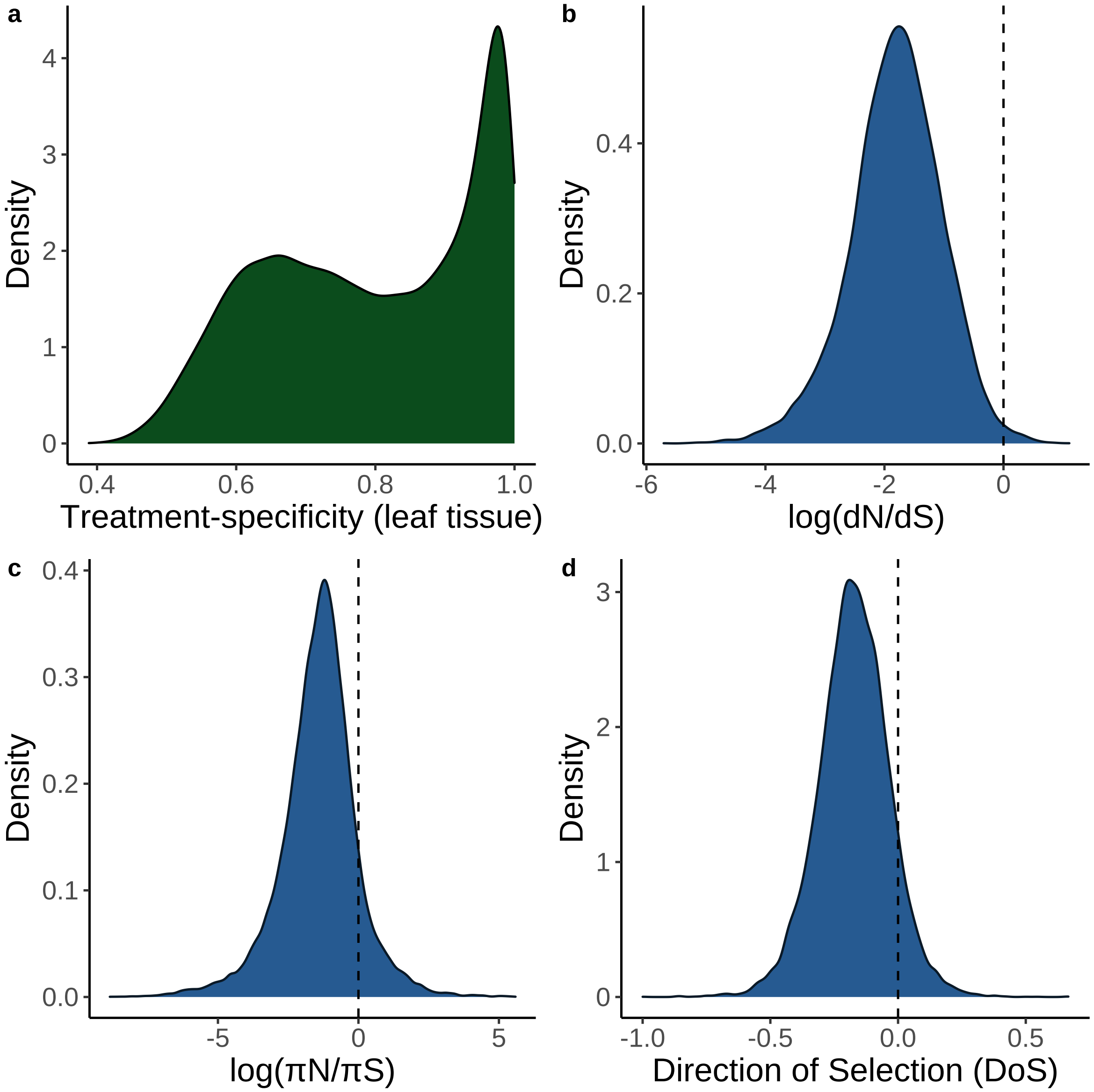

Fig. 2. Density plots of key variables measured in this study. a) Distribution of treatment specificity in leaf tissue expression across all genes included in this study. The area underneath the curve in a given interval of treatment specificity represents the proportion of genes in this study that fall within that range of treatment specificity. b) Distribution of dN/dS across all genes included in this study. The area to the right of the dashed line represents the proportion of genes in this study with $dN/dS > 1$. c) Distribution of π_N/π_S across all genes included in this study. The area to the right of the dashed line represents the proportion of genes in this study with $\pi_N/\pi_S > 1$. d) Distribution of DoS across all genes in this study. Area to the right of the dashed line represents the proportion of genes with DoS > 0, which is interpreted as evidence of adaptive evolution.

expression data on each tissue type and did not account for batch effects (see next section). Among leaf tissue samples, average expression had significant partial correlations with dN ($\rho = -0.19$, $P - \text{value} = 2.1 \times 10^{-122}$) and π_N ($\rho = -0.17$, $P - \text{value} = 2.8 \times 10^{-175}$) after controlling for other factors (Fig. 4, a and b). Treatment specificity was more strongly correlated with dN ($\rho = 0.10$, $P - \text{value} = 7.6 \times 10^{-31}$) and π_N ($\rho = 0.10$, $P - \text{value} = 1.2 \times 10^{-62}$) than Tajima's D ($\rho = 0.03$, $P - \text{value} = 3.1 \times 10^{-7}$) and DoS ($\rho = 0.04$, $P - \text{value} = 2.3 \times 10^{-06}$, Fig. 4, c and d). Furthermore, the top 25% most treatment-specific genes in leaf tissue for our dataset have average dN and π_N values nearly 2.5 times greater than the 25% least treatment-specific genes ($dN = 0.025$ vs 0.061; $\pi_N = 0.0014$ vs 0.0032). Meanwhile, the most and least treatment-specific genes have average Tajima's D values of are −0.44 and −0.43, respectively, and average DoS values of −0.19 and −0.14, respectively. The strongest partial correlation generally occurred between tissue specificity and treatment specificity (Spearman's $\rho = 0.53 - 0.60$, Fig. 4). Gene family size had among the weakest partial correlations with π_N compared to other covariates, but strongly correlated with treatment specificity ($\rho = 0.12$, $P - \text{value} = 6.3 \times 10^{-84}$, Fig. 4b). All of these findings generally held when average expression and treatment specificity were calculated on data from other tissues (Supplementary Table S2, Figs. S6–S10).

Correlations between treatment specificity and nonsynonymous variation persist after controlling for batch effects and dataset imbalance

While combining gene expression data across multiple studies can increase the statistical power of an analysis, there are some

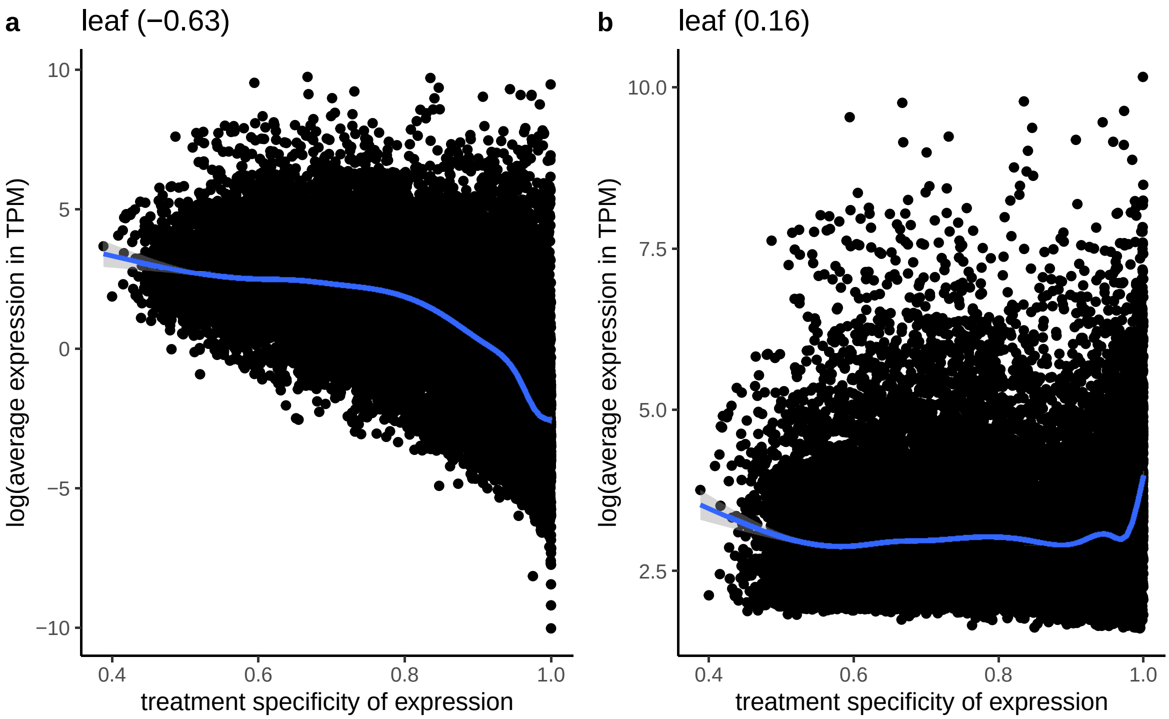

Fig. 3. Correlation between the average expression in transcripts per million (TPM) and treatment specificity of genes when samples with low expression (<5 TPM) are included a) vs excluded b). Expression level and treatment specificity were calculated using only data from leaf tissue samples. Line is a smoothing line with 95 % confidence intervals and values in parentheses give spearman correlation.

potential concerns. First, if many tissue–treatment combinations are not sampled, the dataset will be unbalanced and the effects of tissue and treatment variation on expression could be confounded. Consistent with this expectation, there was a high correlation between tissue specificity and treatment specificity in our initial analyses (Fig. 4, Supplementary Figs. S6–S10). Furthermore, combining data from multiple laboratories could generate batch effects (Leek *et al.* 2010). To address the issues of imbalance and batch effects, we first subset our data to only include treatments where all tissue types were represented. This subset included the treatments of control, abscisic acid, continuous light, warm/hot air temperature, and cold air temperature. We then used SVA to correct for the influence of unknown batch effects on these data subset (Leek and Storey 2007). After SVA, treatment specificity positively correlated with dN ($\rho = 0.10$, $P - \text{value} = 1.6 \times 10^{-32}$) and π_N ($\rho = 0.07$, $P - \text{value} = 1.5 \times 10^{-23}$) when average expression and treatment specificity were calculated on combined fruit and flower data (Supplementary Fig. S33). However, treatment specificity in other tissue types generally did not correlate with our measures of selection (Supplementary Figs. S28–S33, Table S4).

The inclusion of only five treatments in the above analysis could limit quantification of a gene's treatment specificity. Thus, in order to include data from a larger number of treatments, avoid dataset imbalance, and avoid batch effects, we split our expression matrix into six subsets by tissue category. We then further removed treatments that only had expression data from one study to avoid confounding treatment effects with study-specific batch effects. We applied SVA (Leek and Storey 2007) to each of these tissue-specific subsets. After SVA, the expression profiles of most genes appear less treatment-specific (Supplementary Figs. S16–S21 panels a vs b). We also observed less separation in PCA space within treatment groups after SVA (for example, see Supplementary Fig. S16, c and d). Average expression levels before SVA were generally correlated with expression levels after SVA (Supplementary Figs. S16–S21 panels a and b). In partial correlations for theSVA-corrected leaf tissue subset, treatment specificity significantly correlated with dN ($\rho = 0.13$, $P - \text{value} = 6.9 \times 10^{-50}$) and π_N ($\rho = 0.16$, $P - \text{value} = 3.9 \times 10^{-128}$) but less strongly correlated with Tajima's D ($\rho = 0.04$, $P - \text{value} = 6.6 \times 10^{-10}$) and DoS ($\rho = 0.05$, $P - \text{value} = 2.0 \times 10^{-8}$) (Table 1, Fig. 5). These patterns were similar in other tissue types (Supplementary Figs. S11–S15, Table S3).

Discussion

Our main finding is that genes with more treatment-specific expression patterns are, on average, under weaker selective constraint in *A. thaliana*. This is evident by treatment-specific genes generally having higher values of π_N and dN, but not higher values of Tajima's D and DoS, compared to genes with more constitutive expression (Figs. 4 and 5). Our result does not refute the possibility of strong positive selection on treatment-specific genes, as is the case for nucleotide binding site leucine-rich repeat proteins (NBS-LRRs) in *A. thaliana* (Mondragón-Palomino *et al.* 2002). Rather, treatment-specific genes are simply under weaker selection on average compared with less treatment-specific genes. Altogether, this pattern is consistent with the hypothesis that a trade-off between the strength of selection and the treatment specificity of expression helps maintain variation in plasticity for *A. thaliana* (Snell-Rood *et al.* 2010; Van Dyken and Wade 2010).

There are a few ways to think about the biological relevance of the correlations of treatment specificity with π_N and dN. First, the magnitude of treatment specificity's correlation with π_N and dN was generally half the magnitude of average expression's correlation with π_N and dN and similar to tissue specificity's correlation

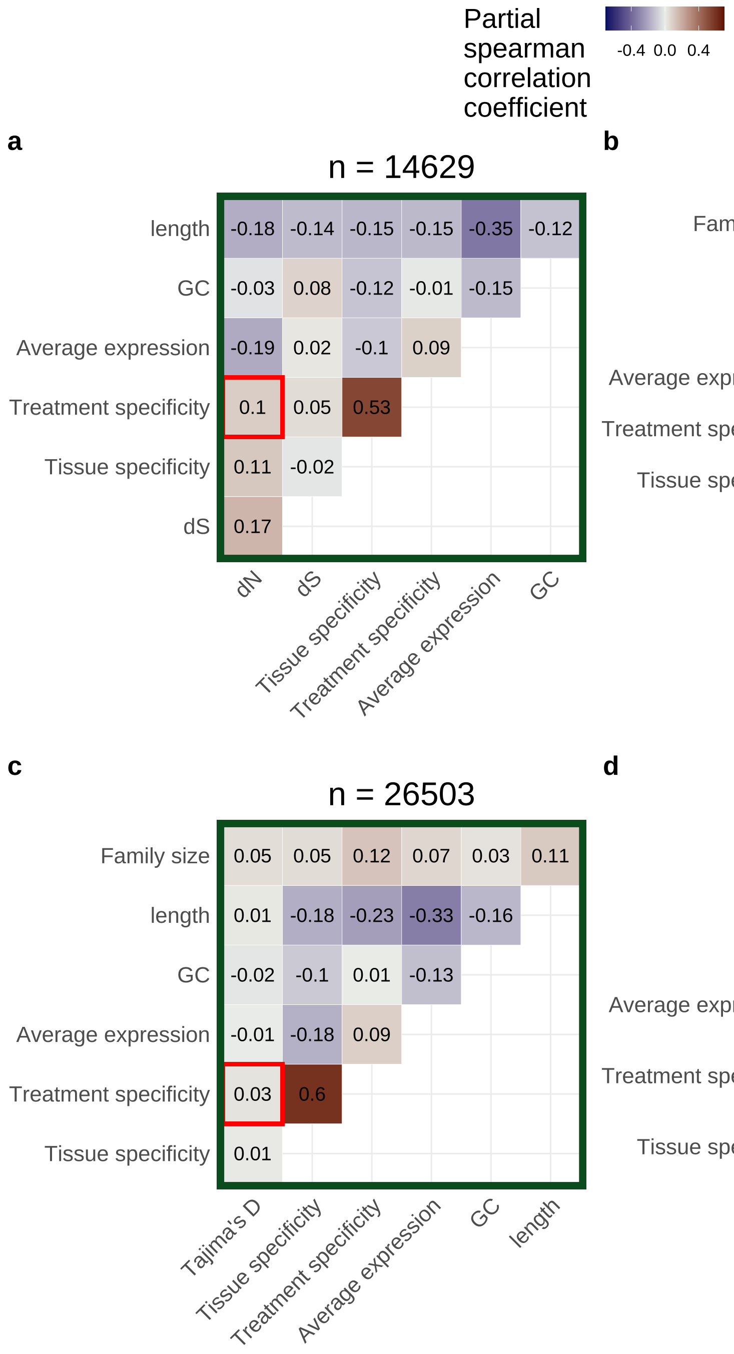

Fig. 4. Partial correlation analysis including either a) *dN*, b) π_N, c) Tajima's D, or d) direction of selection (DoS) as a covariate. Average expression excludes values <5 TPM and was calculated using only leaf tissue samples. Treatment specificity was also calculated using only leaf tissue samples. Tissue specificity was calculated using only control samples across all tissue categories. The number of genes included in each partial correlation analysis (n) is listed at the top of each heatmap.

with π_N and *dN*. Both tissue specificity and average expression are thought to be important determinants of protein evolution (Bush *et al.* 2015; Wu *et al.* 2022), suggesting the comparable effects of treatment specificity may be important too. Second, the effect of treatment specificity on π_N and *dN* persisted even after simultaneously controlling for expression level, tissue specificity, gene length, GC content, and batch effects. Finally, the top 25% most treatment-specific genes in leaf tissue for our dataset have average *dN* and π_N values nearly 2.5 times greater than the 25% least treatment-specific genes ($dN = 0.025$ vs 0.061; $\pi_N = 0.0014$ vs 0.0032), but relatively similar Tajima's D and DoS values (Tajima's D = −0.44 vs −0.43; DoS = −0.19 vs −0.14). These observations together suggest that treatment specificity is an important determinant of protein evolution.

This study disentangles several processes that were often difficult to resolve in previous research. First, many previous studies focus mainly on explaining trends in *dN*/*dS* (Gaut *et al.* 2011; Slotte *et al.* 2011; Bush *et al.* 2015), but both relaxed negative selection and increased positive selection can lead to increases in *dN*/*dS*. To tease apart these two processes, we additionally investigated treatment specificity's relationship with Tajima's D and DoS. Treatment specificity's weaker correlation with Tajima's D

Table 1. Partial correlations between treatment specificity and different measures of selection pre-SVA and post-SVA.

Pre/ post-SVA	Measure of selection	Partial correlation between selection and treatment specificity[a]	*P*-value[b]
Pre	dN	0.10	7.6×10^{-31}
Post	dN	0.13	6.9×10^{-50}
Pre	π_N	0.10	1.2×10^{-62}
Post	π_N	0.16	3.9×10^{-128}
Pre	Tajima's D	0.03	3.1×10^{-7}
Post	Tajima's D	0.04	6.6×10^{-10}
Pre	DoS	0.04	2.3×10^{-6}
Post	DoS	0.05	2.0×10^{-8}

[a] All correlation coefficients are spearman coefficients and are calculated only on leaf tissue samples. [b] All *P*-values represent whether correlation coefficient significantly differs from 0.

and DoS, compared to dN and π_N, suggests that relaxed negative selection plays a larger role than increased positive selection in explaining the high evolutionary rates of treatment-specific genes. Furthermore, measures of expression specificity are often highly correlated with expression level (Slotte *et al.* 2011; Alvarez-Ponce and Fares 2012; Huang 2022). When calculating a gene's expression level, we only included samples where said gene was expressed (TPM > 5) to get an estimate of expression level that was still correlated with dN and π_N, but was independent of expression specificity, allowing us to better disentangle these factors. Finally, previous studies have struggled to partition the factors that influence selection on genes in the presence of predictor variables with considerable error, such as expression level (Drummond *et al.* 2006; Plotkin and Fraser 2007; Yang and Gaut 2011). Error in expression measurements can often be attributed

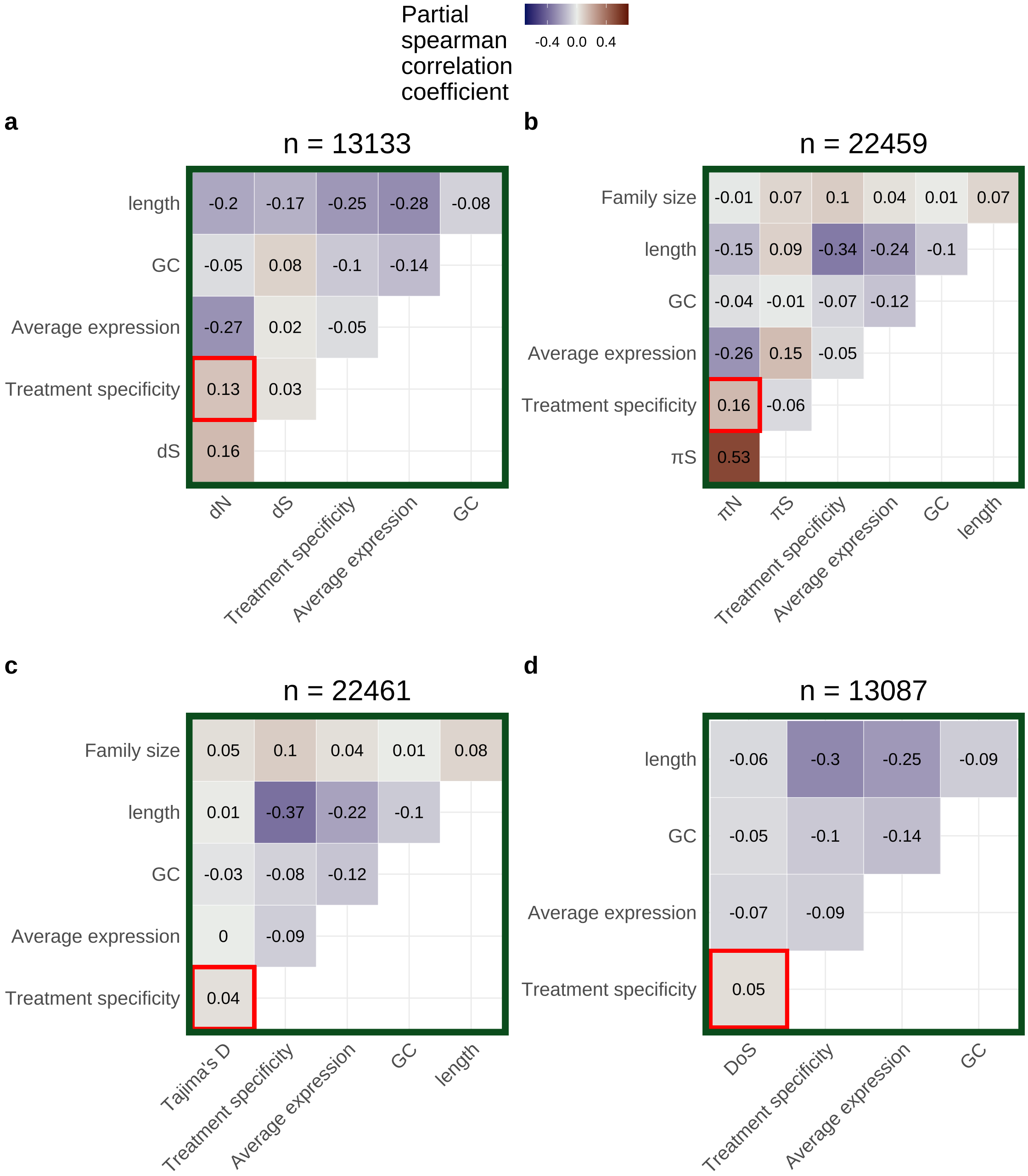

Fig. 5. Partial correlations for a) dN, b) π_N, c) Tajima's D, and d) direction of selection (DoS) based on leaf tissue data subset after applying SVA. Data were further subset to include only treatment groups with data from more than one study before applying SVA. Average expression calculation excludes values <5 TPM. The number of genes included in each partial correlation analysis (n) is listed at the top of each heatmap.

to unmeasured differences between RNA-sequencing experiments (Leek *et al.* 2010) and we accounted for these differences using SVA (Leek and Storey 2007). Even after SVA, treatment specificity was strongly correlated with *dN* and π_N (Fig. 5, a and b), suggesting our results are not an artifact of errors in expression measurement or combining expression data across many studies.

Surprisingly, nearly all genes in *A. thaliana* have some degree of treatment specificity in their expression (Fig. 2a, Supplementary Fig. S3), reflecting results of previous studies on tissue specificity (Eisenberg and Levanon 2003). The high prevalence of treatment specificity in our dataset is partly explained by batch effects because SVA significantly lowered the apparent treatment specificity of most genes (Supplementary Figs. S16b–S21b) and reduced within-treatment differentiation in PCA space (for example, see Supplementary Fig. S16, c and d). This reduction in treatment specificity likely happened because batch effects can include unrecorded between-treatment differences (e.g. the humidity of the growth chamber, light intensity, watering schedule, etc.). Controlling for these unrecorded between-treatment differences thus causes the expression of genes to be less treatment-specific. However, even after batch correction most genes still showed some degree of treatment specificity (Supplementary Fig. S16b–S21b), suggesting it is rare for a gene to be expressed at the same level across many environments.

We also observed that genes with higher treatment specificity generally belonged to larger gene families. We expected gene family size to correlate with selection because singleton and duplicated genes often evolve at different rates (Davis and Petrov 2004; Jordan *et al.* 2004). Theory also suggests that gene duplication relaxes selection on duplicates, allowing for neo- and subfunctionalization (Lynch and Conery 2000; Aagaard *et al.* 2006). We could not investigate how gene family size correlates with *dN* or DoS because measuring these quantities requires identifying substitutions between orthologous genes. Thus, *dN* and *DoS* can only be reliably measured for 1:1 orthologs between *A. thaliana* and *A. lyrata*. However, π_N and Tajima's D can be calculated for genes in larger families and we did observe persistent correlations between family size and Tajima's D (For Fig. 5c: $\rho = 0.05$, $P - \text{value} = 3.1 \times 10^{-12}$; also see Supplementary Figs. S6c–S15c, S28c–S33c). Altogether, these correlations suggest that processes of gene duplication, neofunctionalization, and subfunctionalization could be connected to evolving some degree of treatment specificity.

Gene length was generally the second most correlated factor with *dN* and π_N in our study, just behind average expression. This is consistent with previous work suggesting that longer proteins require more energy to synthesize and are thus under stronger selective constraints (Urrutia and Hurst 2001; Castillo-Davis *et al.* 2002; Eisenberg and Levanon 2003; Urrutia and Hurst 2003). However, while some previous studies in *A. thaliana* observe this same trend (Bush *et al.* 2015), others do not (Slotte *et al.* 2011). This discrepancy could be due to differences in how gene length is defined between studies. In this study, each gene's length included coding sequence as well as introns and untranslated regions, whereas other studies break down gene length into individual features (Bush *et al.* 2015). The goal of this study was not to understand differences in evolution between different gene features, so we included all gene features in our estimate of gene length. However, introns and untranslated regions experience different evolutionary patterns than coding sequences; for example, highly expressed genes being under selection for shorter introns (Castillo-Davis *et al.* 2002; Eisenberg and Levanon 2003). Therefore, future studies must clearly define even seemingly simple features like gene length to ensure that results are comparable across studies.

Although we focused on testing the idea that treatment specificity is responsible for relaxed negative selection in some genes, it is also possible that relaxed selection caused the evolution of treatment specificity. There is some evidence that relaxation of selection occurs before the evolution of expression specificity (Hunt *et al.* 2011) and may better explain cases of neo- and subfunctionalization (Lynch and Conery 2000; Aagaard *et al.* 2006). Future experiments that look at the evolution of treatment specificity and sequence evolution across a broader phylogenetic scale may be helpful for determining the order of these processes.

In summary, this study investigates a trade-off between the treatment-specific expression of a gene and the strength of selection said gene experiences, which is hypothesized to limit plasticity evolution. Consistent with this hypothesis, genes in *A. thaliana* with more treatment-specific expression are under weaker selection compared to more evenly expressed genes. While we find that this trade-off exists, we could not dissect the direction of causality in the trade-off or determine how much this trade-off constrains plasticity evolution relative to other processes. However, these are exciting areas of future research. Future studies should ideally generate fully balanced datasets on gene expression acquired across natural environmental gradients. Taking these steps will contribute to a comprehensive understanding of the constraints on plasticity and protein evolution.

Data availability

All code for our bioinformatic workflows, data analysis, and figure creation can be found here: https://github.com/milesroberts-123/arabidopsis-conditional-expression. The tissue type and treatment annotations for RNA-seq runs in our study can be found in Supplementary Table S5. Genomic references, mapping rates, and a table of expression specificity; nucleotide diversity; and substitution rate values estimated for all *A. thaliana* genes included in this manuscript's analyses is available at: https://doi.org/10.5061/dryad.xd2547dnd. The genome assembly and annotation used in this study was originally downloaded from Phytozome: https://phytozome-next.jgi.doe.gov/. Supplemental material is available at *GENETICS* online.

Acknowledgments

We would like to thank the following individuals for helpful comments on the initial drafts of this manuscript: Robert VanBuren, Shinhan Shiu, Stephen I. Wright, Nathan Catlin, Rebecca Panko, Asia Hightower, Maya Wilson Brown, Sophie Buysse, and Mia Stevens. We would like to further thank the two anonymous peer-reviewers whose comments significantly improved later drafts of this manuscript. Finally, we want to acknowledge the many scientists throughout the world that measured gene expression in *A. thaliana* and made their data publicly available. Without them, this work would not be possible.

Funding

This work was funded by Michigan State University College of Natural Sciences Early Recruitment Fellowship to M.D.R., a National Institutes of Health grant (R35 GM142829) to E.B.J., an Integrated Training Model in Plant And Computational Sciences Fellowship (National Science Foundation: DGE-1828149) to M.D.R., a Plant Biotechnology for Health and Sustainability

Fellowship (National Institutes of Health: T32-GM110523) to M.D.R., and a Michigan State University Institute for Cyber-Enabled Research Cloud Computing Fellowship to M.D.R.

Conflicts of interest

The authors declare no conflict of interest.

Literature cited

Aagaard JE, Willis JH, Phillips PC. Relaxed selection among duplicate floral regulatory genes in Lamiales. J Mol Evol. 2006;63:493–503. doi:10.1007/s00239-005-0306-x

Alonso-Blanco C, Andrade J, Becker C, Bemm F, Bergelson J, Borgwardt KM, Cao J, Chae E, Dezwaan TM, Ding W, *et al.* 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. Cell. 2016;166:481–491. doi:10.1016/j.cell.2016.05.063

Alvarez-Ponce D, Fares MA. Evolutionary rate and duplicability in the *Arabidopsis thaliana* protein–protein interaction network. Genome Biol Evol. 2012;4:1263–1274. doi:10.1093/gbe/evs101

Alvarez-Ponce D, Feyertag F, Chakraborty S. Position matters: network centrality considerably impacts rates of protein evolution in the human protein–protein interaction network. Genome Biol Evol. 2017;9:1742–1756. doi:10.1093/gbe/evx117

Auld JR, Agrawal AA, Relyea RA. Re-evaluating the costs and limits of adaptive phenotypic plasticity. Proc R Soc B: Biol Sci. 2010;277: 503–511. doi:10.1098/rspb.2009.1355

Betancourt AJ, Presgraves DC. Linkage limits the power of natural selection in *Drosophila*. Proc Natl Acad Sci USA. 2002;99:13616–13620. doi:10.1073/pnas.212277199

Bush SJ, Kover PX, Urrutia AO. Lineage-specific sequence evolution and exon edge conservation partially explain the relationship between evolutionary rate and expression level in *A. thaliana*. Mol Ecol. 2015;24:3093–3106. doi:10.1111/mec.13221

Castillo-Davis CI, Mekhedov SL, Hartl DL, Koonin EV, Kondrashov FA. Selection for short introns in highly expressed genes. Nat Genet. 2002;31:415–418. doi:10.1038/ng940

Charlesworth B. The effect of background selection against deleterious mutations on weakly selected, linked variants. Genet Res. 1994;63:213–227. doi:10.1017/S0016672300032365

Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34:i884–i890. doi:10.1093/bioinformatics/bty560

Cheng CY, Krishnakumar V, Chan AP, Thibaud-Nissen F, Schobel S, Town CD. Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. Plant J. 2017;89:789–804. doi:10.1111/tpj.13415

Cingolani P, Patel VM, Coon M, Nguyen T, Land SJ, Ruden DM, Lu X. Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift. Front Genet. 2012a;3:35. doi:10.3389/fgene.2012.00035

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly. 2012b;6: 80–92. doi:10.4161/fly.19695

Colombo M, Laayouni H, Invergo BM, Bertranpetit J, Montanucci L. Metabolic flux is a determinant of the evolutionary rates of enzyme-encoding genes. Evolution. 2014;68:605–613. doi:pdf/10.1111/evo.12262

Cooper L, Meier A, Laporte MA, Elser JL, Mungall C, Sinn BT, Cavaliere D, Carbon S, Dunn NA, Smith B, *et al.* The Planteome database: an integrated resource for reference ontologies, plant genomics and phenomics. Nucleic Acids Res. 2018;46:D1168–D1180. doi:10.1093/nar/gkx1152

Crameri F. Geodynamic diagnostics, scientific visualisation and StagLab 3.0. Geosci Model Dev. 2018;11:2541–2562. doi:10.5194/gmd-11-2541-2018

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, *et al.* The variant call format and VCFtools. Bioinformatics. 2011;27:2156–2158. doi:10.1093/bioinformatics/btr330

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, *et al.* Twelve years of SAMtools and BCFtools. GigaScience. 2021;10:giab008. doi:10.1093/gigascience/giab008

Davis JC, Petrov DA. Preferential duplication of conserved proteins in eukaryotic genomes. PLoS Biol. 2004;2:e55. doi:10.1371/journal.pbio.0020055

DeWitt TJ, Sih A, Wilson DS. Costs and limits of phenotypic plasticity. Trends Ecol Evol. 1998;13:77–81. doi:10.1016/S0169-5347(97)01274-3

Drummond DA, Raval A, Wilke CO. A single determinant dominates the rate of yeast protein evolution. Mol Biol Evol. 2006;23:327–337. doi:10.1093/molbev/msj038

Duret L, Mouchiroud D. Determinants of substitution rates in mammalian genes: expression pattern affects selection intensity but not mutation rate. Mol Biol Evol. 2000;17:68–070. doi:10.1093/oxfordjournals.molbev.a026239

Eisenberg E, Levanon EY. Human housekeeping genes are compact. Trends Genet. 2003;19:362–365. doi:10.1016/S0168-9525(03)00140-9

Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019;20:238. doi:10.1186/s13059-019-1832-y

Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32:3047–3048. doi:10.1093/bioinformatics/btw354

Fukushima K, Pollock DD. Amalgamated cross-species transcriptomes reveal organ-specific propensity in gene expression evolution. Nat Commun. 2020;11:4459. doi:10.1038/s41467-020-18090-8

Gaut B, Yang L, Takuno S, Eguiarte LE. The patterns and causes of variation in plant nucleotide substitution rates. Annu Rev Ecol Evol Syst. 2011;42:245–266. doi:10.1146/annurev-ecolsys-102710-145119

Glémin S. Mating systems and the efficacy of selection at the molecular level. Genetics. 2007;177:905–916.

Groen SC, Ćalić I, Joly-Lopez Z, Platts AE, Choi JY, Natividad M, Dorph K, Mauck WM, Bracken B, Cabral CLU, *et al.* The strength and pattern of natural selection on gene expression in rice. Nature. 2020; 578:572–576. doi:10.1038/s41586-020-1997-2

Huang YF. Dissecting genomic determinants of positive selection with an evolution-guided regression model. Mol Biol Evol. 2022; 39:msab291. doi:10.1093/molbev/msab291

Huber CD, Kim BY, Marsden CD, Lohmueller KE. Determining the factors driving selective effects of new nonsynonymous mutations. Proc Natl Acad Sci USA. 2017;114:4465–4470. doi:10.1073/pnas.1619508114

Hunt BG, Ometto L, Wurm Y, Shoemaker D, Yi SV, Keller L, Goodisman MAD. Relaxed selection is a precursor to the evolution of phenotypic plasticity. Proc Natl Acad Sci USA. 2011;108: 15936–15941. doi:10.1073/pnas.1104825108

Jordan IK, Wolf YI, Koonin EV. Duplicated genes evolve slower than singletons despite the initial rate increase. BMC Evol Biol. 2004; 4:22. doi:10.1186/1471-2148-4-22

Josephs EB, Wright SI, Stinchcombe JR, Schoen DJ. The relationship between selection, network connectivity, and regulatory variation within a population of *Capsella grandiflora*. Genome Biol Evol. 2017;9:1099–1109. doi:10.1093/gbe/evx068

Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–780. doi:10.1093/molbev/mst010

Kawecki TJ. Accumulation of deleterious mutations and the evolutionary cost of being a generalist. Am Nat. 1994;144:833–838. doi:10.1086/285709

Kim S. ppcor: an R package for a fast calculation to semi-partial correlation coefficients. Commun Stat Appl Methods. 2015;22: 665–674.

Koonin EV. Are there laws of genome evolution?. PLoS Comput Biol. 2011;7:e1002173. doi:10.1371/journal.pcbi.1002173

Korunes KL, Samuk K. pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. Mol Ecol Resour. 2021;21:1359–1368. doi:10.1111/1755-0998.13326

Kryuchkova-Mostacci N, Robinson-Rechavi M. A benchmark of gene expression tissue-specificity metrics. Brief Bioinformatics. 2017; 18:205–214.

Lanfear R, Ho SYW, Jonathan Davies T, Moles AT, Aarssen L, Swenson NG, Warman L, Zanne AE, Allen AP. Taller plants have lower rates of molecular evolution. Nat Commun. 2013;4: 1879. doi:10.1038/ncomms2836

Langley CH, Stevens K, Cardeno C, Lee YCG, Schrider DR, Pool JE, Langley SA, Suarez C, Corbett-Detig RB, Kolaczkowski B, *et al.* Genomic variation in natural populations of *Drosophila melanogaster*. Genetics. 2012;192:533–598. doi:10.1534/genetics.112.142018

Larracuente AM, Sackton TB, Greenberg AJ, Wong A, Singh ND, Sturgill D, Zhang Y, Oliver B, Clark AG. Evolution of protein-coding genes in *Drosophila*. Trends Genet. 2008;24:114–123. doi: 10.1016/j.tig.2007.12.001

Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE, Geman D, Baggerly K, Irizarry RA. Tackling the widespread and critical impact of batch effects in high-throughput data. Nat Rev Genet. 2010;11:733–739. doi:10.1038/nrg2825

Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. PLoS Genet. 2007;3:e161. doi: 10.1371/journal.pgen.0030161

Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009;25: 1754–1760. doi:10.1093/bioinformatics/btp324

Lin YS, Hsu WL, Hwang JK, Li WH. Proportion of solvent-exposed amino acids in a protein and rate of protein evolution. Mol Biol Evol. 2007;24:1005–1011. doi:10.1093/molbev/msm019

Lynch M, Conery JS. The evolutionary fate and consequences of duplicate genes. Science. 2000;290:1151–1155. doi:10.1126/science. 290.5494.1151

Mähler N, Wang J, Terebieniec BK, Ingvarsson PK, Street NR, Hvidsten TR. Gene co-expression network connectivity is an important determinant of selective constraint. PLoS Genet. 2017; 13:e1006402.

Masalia RR, Bewick AJ, Burke JM. Connectivity in gene coexpression networks negatively correlates with rates of molecular evolution in flowering plants. PLoS ONE. 2017;12:e0182289. doi:10.1371/ journal.pone.0182289

McGuigan K, Collet JM, Allen SL, Chenoweth SF, Blows MW. Pleiotropic mutations are subject to strong stabilizing selection. Genetics. 2014;197:1051–1062. doi:10.1534/genetics.114.165720

Mondragón-Palomino M, Meyers BC, Michelmore RW, Gaut BS. Patterns of positive selection in the complete NBS-LRR gene family of *Arabidopsis thaliana*. Genome Res. 2002;12:1305–1315.

Moutinho AF, Eyre-Walker A, Dutheil JY. Strong evidence for the adaptive walk model of gene evolution in *Drosophila* and *Arabidopsis*. PLoS Biol. 2022;20:e3001775. doi:10.1371/journal. pbio.3001775

Mukherjee D, Mukherjee A, Ghosh TC. Evolutionary rate heterogeneity of primary and secondary metabolic pathway genes in *Arabidopsis thaliana*. Genome Biol Evol. 2016;8:17–28. doi:10. 1093/gbe/evv217

Murren CJ, Auld JR, Callahan H, Ghalambor CK, Handelsman CA, Heskel MA, Kingsolver JG, Maclean HJ, Masel J, Maughan H, *et al.* Constraints on the evolution of phenotypic plasticity: limits and costs of phenotype and plasticity. Heredity. 2015;115: 293–301. doi:10.1038/hdy.2015.8

Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol. 1986;3:418–426.

Nembaware V, Crum K, Kelso J, Seoighe C. Impact of the presence of paralogs on sequence divergence in a set of mouse-human orthologs. Genome Res. 2002;12:1370–1376. doi:10.1101/gr.270902

Nordborg M, Hu TT, Ishino Y, Jhaveri J, Toomajian C, Zheng H, Bakker E, Calabrese P, Gladstone J, Goyal R, *et al.* The pattern of polymorphism in *Arabidopsis thaliana*. PLoS Biol. 2005;3:e196. doi:10. 1371/journal.pbio.0030196

Ossowski S, Schneeberger K, Lucas-Lledó JI, Warthmann N, Clark RM, Shaw RG, Weigel D, Lynch M. The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. Science. 2010;327:92–94. doi:10.1126/science.1180677

Paape T, Bataillon T, Zhou P, Briskine R, Young ND, Tiffin P. Selection, genome-wide fitness effects and evolutionary rates in the model legume *Medicago truncatula*. Mol Ecol. 2013;22:3525–3538. doi:10. 1111/mec.12329

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14:417–419. doi:10.1038/nmeth.4197

Payne BL, Alvarez-Ponce D. Higher rates of protein evolution in the self-fertilizing plant *Arabidopsis thaliana* than in the out-crossers *Arabidopsis lyrata* and *Arabidopsis halleri*. Genome Biol Evol. 2018; 10:895–900. doi:10.1093/gbe/evy053

Pedersen TL, Crameri F. scico: Colour Palettes Based on the Scientific Colour-Maps. R; 2022.

Pigliucci M, Kolodynska A. Phenotypic plasticity to light intensity in *Arabidopsis thaliana*: invariance of reaction norms and phenotypic integration. Evol Ecol. 2002;16:27–47. doi:10.1023/A:1016073525567

Plotkin JB, Fraser HB. Assessing the determinants of evolutionary rates in the presence of noise. Mol Biol Evol. 2007;24:1113–1121. doi:10.1093/molbev/msm044

Rocha EPC. The quest for the universals of protein evolution. Trends Genet. 2006;22:412–416. doi:10.1016/j.tig.2006.06.004

Scheiner SM. Genetics and evolution of phenotypic plasticity. Annu Rev Ecol Syst. 1993;24:35–68. doi:10.1146/annurev.es.24.110193. 000343

Schlichting CD, Smith H. Phenotypic plasticity: linking molecular mechanisms with evolutionary outcomes. Evol Ecol. 2002;16: 189–211. doi:10.1023/A:1019624425971

Schneider H. Characterization, costs, cues, and future perspectives of phenotypic plasticity. Ann Bot. 2022;130:131–148. doi:10. 1093/aob/mcac087

Slotte T, Bataillon T, Hansen TT, Wright SI, Schierup MH. Genomic determinants of protein evolution and polymorphism in *Arabidopsis*. Genome Biol Evol. 2011;3:1210–1219. doi:10.1093/gbe/evr094

Snell-Rood EC, Van Dyken JD, Cruickshank T, Wade MJ, Moczek AP. Toward a population genetic framework of developmental evolution: the costs, limits, and consequences of phenotypic plasticity. BioEssays. 2010;32:71–81. doi:10.1002/bies.200900132

Stoletzki N, Eyre-Walker A. Estimation of the neutrality index. Mol Biol Evol. 2011;28:63–70. doi:10.1093/molbev/msq249

Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 2006;34:W609–W612. doi:10.1093/nar/gkl315

Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123:585–595. doi:10.1093/genetics/123.3.585

Takuno S, Gaut BS. Body-methylated genes in *Arabidopsis thaliana* are functionally important and evolve slowly. Mol Biol Evol. 2012;29: 219–227. doi:10.1093/molbev/msr188

Urrutia AO, Hurst LD. Codon usage bias covaries with expression breadth and the rate of synonymous evolution in humans, but this is not evidence for selection. Genetics. 2001;159:1191–1199. doi:10.1093/genetics/159.3.1191

Urrutia AO, Hurst LD. The signature of selection mediated by expression on human genes. Genome Res. 2003;13:2260–2264. doi:10.1101/gr.641103

Van Buskirk J, Steiner UK. The fitness costs of developmental canalization and plasticity. J Evol Biol. 2009;22:852–860. doi:10.1111/j.1420-9101.2009.01685.x

Van Dyken JD, Wade MJ. The genetic signature of conditional expression. Genetics. 2010;184:557–570. doi:10.1534/genetics.109.110163

Van Kleunen M, Fischer M. Constraints on the evolution of adaptive phenotypic plasticity in plants. New Phytol. 2005;166:49–60. doi:10.1111/j.1469-8137.2004.01296.x

Wang Hc, Singer GAC, Hickey DA. Mutational bias affects protein evolution in flowering plants. Mol Biol Evol. 2004;21:90–96. doi:10.1093/molbev/msh003

Wheeler LC, Walker JF, Ng J, Deanna R, Dunbar-Wallis A, Backes A, Pezzi PH, Palchetti MV, Robertson HM, Monaghan A, *et al.* Transcription factors evolve faster than their structural gene targets in the flavonoid pigment pathway. Mol Biol Evol. 2022;39: msac044. doi:10.1093/molbev/msac044

Whitlock MC. The red queen beats the jack-of-all-trades: the limitations on the evolution of phenotypic plasticity and niche breadth. Am Nat. 1996;148:S65–S77. doi:10.1086/285902

Winter EE, Goodstadt L, Ponting CP. Elevated rates of protein secretion, evolution, and disease among tissue-specific genes. Genome Res. 2004;14:54–61. doi:10.1101/gr.1924004

Wright SI, Lauga B, Charlesworth D. Rates and patterns of molecular evolution in inbred and outbred *Arabidopsis*. Mol Biol Evol. 2002; 19:1407–1420. doi:10.1093/oxfordjournals.molbev.a004204

Wright SI, Yau CBK, Looseley M, Meyers BC. Effects of gene expression on molecular evolution in *Arabidopsis thaliana* and *Arabidopsis lyrata*. Mol Biol Evol. 2004;21:1719–1726. doi:10.1093/molbev/msh191

Wu Z, Cai X, Zhang X, Liu Y, Tian Gb, Yang JR, Chen X. Expression level is a major modifier of the fitness landscape of a protein coding gene. Nat Ecol & Evol. 2022;6:103–115.

Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, Ophir R, Bar-Even A, Horn-Saban S, Safran M, Domany E, *et al.* Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. Bioinformatics. 2005;21:650–659. doi:10.1093/bioinformatics/bti042

Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007;24:1586–1591. doi:10.1093/molbev/msm088

Yang L, Gaut BS. Factors that contribute to variation in evolutionary rate among *Arabidopsis* genes. Mol Biol Evol. 2011;28:2359–2369. doi:10.1093/molbev/msr058

Yang J, Gu Z, Li WH. Rate of protein evolution versus fitness effect of gene deletion. Mol Biol Evol. 2003;20:772–774. doi:10.1093/molbev/msg078

Zhang L, Li WH. Mammalian housekeeping genes evolve more slowly than tissue-specific genes. Mol Biol Evol. 2004;21:236–239. doi:10.1093/molbev/msh010

Zhang L, Vision TJ, Gaut BS. Patterns of nucleotide substitution among simultaneously duplicated gene pairs in *Arabidopsis thaliana*. Mol Biol Evol. 2002;19:1464–1473. doi:10.1093/oxfordjournals.molbev.a004209

Zhang J, Yang JR. Determinants of the rate of protein sequence evolution. Nat Rev Genet. 2015;16:409–420. doi:10.1038/nrg3950

Zhang H, Zhang F, Yu Y, Feng L, Jia J, Liu B, Li B, Guo H, Zhai J. A comprehensive online database for exploring 20,000 public *Arabidopsis* RNA-seq libraries. Mol Plant. 2020;13:1231–1233. doi:10.1016/j.molp.2020.08.001

Editor: D. Begun