# 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**
**2 Department of Plant Biology, Michigan State University, East Lansing, MI**
**3 Ecology, Evolution, and Behavior Program, Michigan State University, East Lansing, MI**

*** milesdroberts@gmail.com, 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.

## 1 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 [64]. Plastic responses are partly controlled through differential gene expression between environments [67, 68]. 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 [80]. These constraints are usually characterized as either costs, where plasticity reduces fitness in some way, or limits to the evolution or maintenance of plasticity [20]. Decades of research has attempted to measure the costs associated with plasticity (reviewed in [69]) but studies often fail to detect costs or find costs that are weak or restricted to certain environments

[80, 78, 5]. Theory also predicts that there will be strong selection to alleviate costs [54]. Thus, limits may be more important than costs in shaping the evolution of plasticity.

Recent work suggests that relaxed selection can limit plasticity evolution [71, 54]. For instance, one hypothesis posits that genes are often under selection for environment-specific expression to minimize deleterious pleiotropy [71, 50, 31]. 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 [36, 83, 79]. 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 [36, 71]. 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 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 [66, 27, 38, 93]) and these factors are hard to disentangle. A protein's expression level is often considered the best predictor of its evolutionary rate [66] - a result observed across all domains of life [93] and sometimes considered a "law" of genome evolution [38]. Among multicellular organisms, the degree of tissue specificity in expression is also generally predictive of evolutionary rates [22, 43, 84, 94, 70, 7, 53, 29, 30]. Additional factors that also influence evolutionary rates include exon edge conservation [7], mutational bias [81, 59], gene length [53], gene age [52], GC content [95, 53], expression stochasticity [29], involvement in general vs specialized metabolism [53], identity as a regulatory or structural gene [82], recombination rate [42], codon-bias [6], mating system [85, 28, 62], gene compactness [43, 53], co-expression or protein-protein interaction network connectivity [3, 49, 55, 4, 34], gene body methylation [75], metabolic flux [14], protein structure [47], essentiallity [57, 89, 19], and even plant height [41]. 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.

# 2 Materials and methods

## 2.1 RNA-seq run annotation

We amassed an initial set of RNA-seq runs from the Sustech Arabidopsis RNA-seq database V2 [92] (`http://ipf.sustech.edu.cn/pub/athrdb/`) excluding any samples not associated with a publication or lacking a tissue type label. On May 24th, 2022 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) [15], respectively, based on information in their associated publications and SRA metadata. In our analysis, control exposure was defined as long day conditions (12 hrs 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.

## 2.2 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 [10], 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 [25]. All trimmed runs were then aligned to a decoy-aware transcriptome index made by combining the primary transcripts of the Araport11 genome annotation [11] with the *A. thaliana* genome in salmon v1.2.1 [61] using an index size of 25bp. 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.

## 2.3 Whole genome sequence data processing

We downloaded whole genome sequencing data for 1135 *A. thaliana* accessions from the 1001 genomes project panel (SRA project SRP056687) [2] using the SRA toolkit. All

runs were trimmed using fastp [10], 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 [46]. The alignments were sorted and converted to BAM format with SAMTOOLS v1.11 [18], 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 [39], variant and invariant sites were filtered separately using both GATK and vcftools v0.1.15 [17]. 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 2 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) [13] and variants labeled as either missense or synonymous were separated into different files using SnpSift [12].

## 2.4 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 [24]. For each 1:1 ortholog, we aligned their protein sequences with MAFFT L-INS-I v7.475 [35], then converted the protein alignments to gapless codon-based alignments using pal2nal v14 [73]. Using the gapless codon-based alignments, we estimated $dN/dS$ using the method in [56] implemented as a custom Biopython v1.79 script and implemented through the codeml program in the PAML package v4.9 [91]. Unlike codeml, the custom Biopython script also returns counts of nonsynonymous ($N$) and synonymous sites ($S$) within each gene as described in [56], which we later used to calculate nucleotide diversity per nonsynonymous site ($\pi_N$) and per synonymous site ($\pi_S$). Before proceeding with more analyses, we confirmed that our estimates of $dN$ and $dS$ were consistent between our Biopython script and codeml (Figure S5, Pearson correlations $dN : \rho = 0.9998$, $dS : \rho = 0.9809$). The outputs of the Biopython script were used in all subsequent analyses.

## 2.5 Selection estimated from within-species polymorphism

### 2.5.1 Nucleotide diversity at nonsynonymous sites.

Nucleotide diversity ($\pi$) was calculated for each gene with pixy v1.2.3.beta1 [39] 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 $\pi$, $\pi_N$, and $\pi_S$, respectively, by first multiplying the per site estimate output from pixy by the number of sites included in the analysis. Then, to get $\pi_N$ and $\pi_S$, 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 [56].

### 2.5.2 Tajima's D.

We next calculated Tajima's D for each gene. First, we calculated $\pi$ and Watterson's Theta ($\theta_W$) for each variant site $i$ within a gene ($\pi_i$ and $\theta_{Wi}$ respectively). In this case, $\pi_i$ 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, $\theta_{Wi}$ 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 $\theta_{Wi}$ is equivalent to the usual calculation of $\theta_W$ with the number of segregating sites set to one. Next, the variance in Tajima's D was calculated for each site as:

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

This is equivalent to equation 38 in [74] with the number of segregating sites set to one.

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

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

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

### 2.5.3 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) [72] as follows:

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

We chose this metric, as opposed to the proportion of amino acid substitutions driven by positive selection ($\alpha$), because it is less biased than $\alpha$ [72] and was successfully used in studies similar to ours [60]. Furthermore, we found that $\alpha$ often returns uninterpretable negative values when applied to *A. thaliana*, perhaps because of an excess of slightly deleterious polymorphisms [58] due to their predominantly selfing mating system [9].

## 2.6 Treatment specificity

Treatment specificity ($\tau$) was estimated separately for runs from each tissue type using the following formula [88]:

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

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 $N$ means that $\tau$ 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 [40] and is normalized by the number of treatments included, making it comparable across data sets. We also applied the same formula to calculate tissue specificity in several different treatment conditions.

## 2.7 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, $\tau$. In both simulations, we generated 1000 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 5 groups of columns, with 5 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 $\tau$ 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.

## 2.8 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 [70]. 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 [8, 23]. Finally, the family size for each gene was estimated as the number of *A. thaliana* genes in their respective orthogroups output by OrthoFinder.

## 2.9 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 ($dN$, $\pi_N$, Tajima's D, and $DoS$) after accounting for average expression (excluding values TPM ¡ 5), gene length, and GC content using the ppcor R package [37]. For partial correlation analyses involving $\pi_N$ 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 [16, 63].

## 2.10 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 [26]. 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 [45] 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 $j$, $\mu_i$ 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

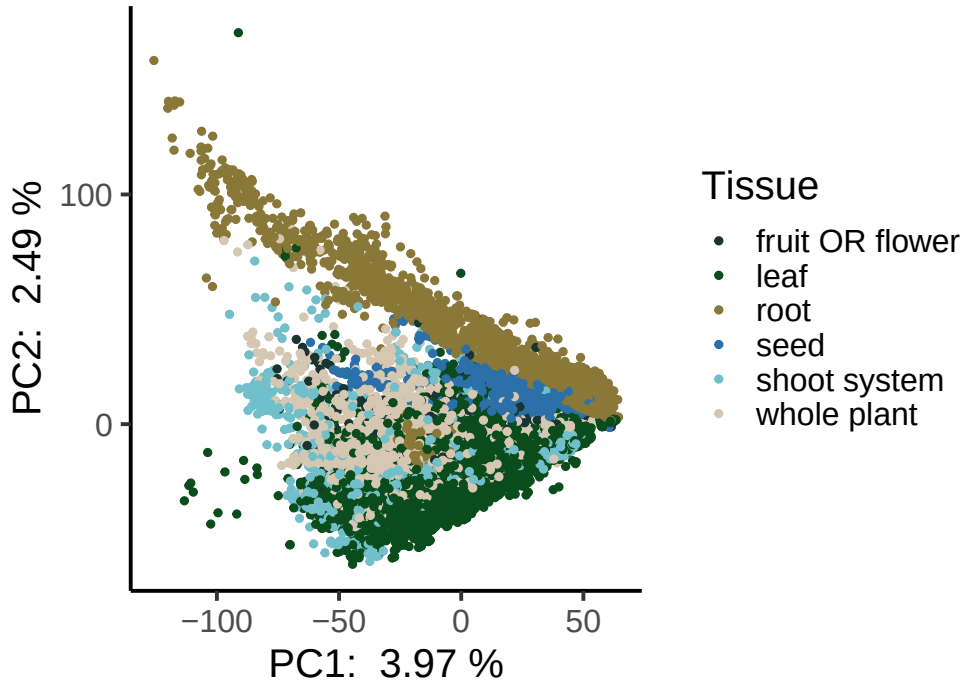


**Figure 1.** Principal components analysis of all expression data. Each point represents a different RNA-seq experiment and is colored by its associated tissue type. Experiments from all treatment conditions are included in this analysis. Plotting order was randomized to avoid overplotting.

estimate what the unmodeled variables influencing expression are, but rather find a set of vectors (the surrogate variables) that span the same space as $\mathbf{g}$:

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

Where each $\mathbf{h_k}$ is a surrogate variable and each $\lambda_\mathbf{k}$ 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.

# 3 Results

## 3.1 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 Tables 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 % (Figure 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 (Figure 1). We also observed that nearly all genes had some degree of treatment specificity in their expression (Figures 2A, S3). Furthermore, only a small proportion of genes had strong signatures of selection based on $dN/dS$, $\pi_N/\pi_S$, DoS, or Tajima's D (Figure 2B-D, Figure S2). The treatment specificity of expression was lower on average in flower and fruit tissue compared to the other tissues (Figure S3). However, tissue specificity did not vary widely depending on the treatment condition (Figure S4).

## 3.2 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 [86]. 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 (Figure 3) and other tissues (Figures S34 - S38) and replicated the result by simulating gene expression matrices (Figure 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).

## 3.3 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

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-value $= 2.1 \times 10^{-122}$) and $\pi_N$ ($\rho = -0.17$, p-value $= 2.8 \times 10^{-175}$) after controlling for other factors (Figures 4A,4B). Treatment specificity was more strongly correlated with $dN$ ($\rho = 0.10$, p-value $= 7.6 \times 10^{-31}$) and $\pi_N$ ($\rho = 0.10$, p-value $= 1.2 \times 10^{-62}$) than Tajima's D ($\rho = 0.01$, p-value $= 3.1 \times 10^{-7}$) and DoS ($\rho = 0.04$, p-value $= 2.3 \times 10^{-06}$, Figure 4C, 4D). Furthermore, the top 25% most treatment-specific genes in leaf tissue for our dataset have average $dN$ and $\pi_N$ 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 $\rho = 0.53 - 0.60$, Figure 4). Gene family size had among the weakest partial correlations with $\pi_N$ compared to other covariates, but strongly correlated with treatment specificity ($\rho = 0.12$, p-value $= 6.3 \times 10^{-84}$, Figure 4B). All of these findings generally held when average expression and treatment specificity were calculated on data from other tissues (Table S2, Figures S6-S10).

### 3.4 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 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 (Figure 4, S6-S10). Furthermore, combining data from multiple laboratories could generate batch effects [44]. 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 this data subset [45]. After SVA, treatment specificity positively correlated with $dN$ ($\rho = 0.10$, p-value $= 1.6 \times 10^{-32}$) and $\pi_N$ ($\rho = 0.07$, p-value $= 1.5 \times 10^{-23}$) when average expression and treatment specificity were calculated on combined fruit and flower data (Figures S33). However, treatment specificity in other tissue types generally did not correlate with our measures of selection (Figures 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 [45] to each of these tissue-specific subsets. After SVA, the expression profiles of most genes appear less treatment-specific (Figures S16-S21 panels A vs B). We also observed less separation in PCA space within treatment groups after SVA (for example, see Figures S16C and S16D). Average expression levels before SVA were generally correlated with expression levels after SVA (Figures S16-S21 panels A and B). In partial correlations on each SVA-corrected subset, treatment specificity significantly correlated with $dN$ ($\rho = 0.13$, p-value $= 6.9 \times 10^{-50}$) and $\pi_N$ ($\rho = 0.16$, p-value $= 3.9 \times 10^{-128}$) but less strongly correlated with Tajima's D

($\rho = 0.04$, p-value $= 6.6 \times 10^{-10}$) and DoS ($\rho = 0.05$, p-value $= 2.0 \times 10^{-8}$) for the leaf tissue data subset (Table 1, Figures 5). These patterns were similar in other tissue types (Figures S11-S15, Table S3).

# 4 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 $\pi_N$ and $dN$, but not higher values of Tajima's D and DoS, compared to genes with more constitutive expression (Figures 4,5). 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* [51]. Rather, treatment-specific genes are simply under weaker selection on average compared to 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* [71, 79].

There are a few ways to think about the biological relevance of the correlations of treatment specificity with $\pi_N$ and $dN$. First, the magnitude of treatment specificity's correlation with $\pi_N$ and $dN$ was generally half the magnitude of average expression's correlation with $\pi_N$ and $dN$ and similar to tissue specificity's correlation with $\pi_N$ and $dN$. Both tissue specificity and average expression are thought to be important determinants of protein evolution [7, 87], suggesting the comparable effects of treatment specificity may be important too. Second, the effect of treatment specificity on $\pi_N$ 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 our dataset have average $dN$ and $\pi_N$ 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 $dN/dS$ [70, 27, 7], 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 and DoS, compared to $dN$ and $\pi_N$, 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 [70, 3, 30]. 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 $\pi_N$, 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 [21, 65, 90]. Error in expression measurements can often be attributed to unmeasured differences between RNA-sequencing experiments [44] and we accounted for these differences using SVA [45]. Even after SVA, treatment specificity was strongly correlated with $dN$ and $\pi_N$ (Figures 5A-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 (Figures 2A, S3), reflecting results of previous studies on tissue specificity [23]. 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 (Figures S16B-S21B) and reduced within-treatment differentiation in PCA space (for example, see Figures S16C and S16D). 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 (Figures 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 [33, 19]. Theory also suggests that gene duplication leads to relaxation of selection on duplicates, allowing for neo- and sub-functionalization [48, 1]. 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, $\pi_N$ 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 Figure 5C: $\rho = 0.05$, p-value $= 3.1 \times 10^{-12}$; also see Figures 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 $\pi_N$ 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 [76, 8, 23, 77]. However, while some previous studies in *A. thaliana* observe this same trend [7], others do not [70]. 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 [7]. 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 [8, 23]. 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 [32] and may better explain cases of neo- and subfunctionalization [48, 1]. 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.

# 5 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 Table S5. Genomic references as well as 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: XXX (url inserted at publication). The genome assembly and annotation used in this study was originally downloaded from Phytozome: `https://phytozome-next.jgi.doe.gov/`.

# 6 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 hundreds of 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.

# 7 Funding

This work was funded by Michigan State University's College of Natural Sciences Early Recruitment Fellowship to MDR, a National Institutes of Health grant (R35 GM142829) to EBJ, an Integrated Training Model in Plant And Computational Sciences Fellowship (National Science Foundation: DGE-1828149) to MDR, a Plant Biotechnology for Health and Sustainability Fellowship (National Institutes of Health: T32-GM110523) to MDR, and a Michigan State University Institute for Cyber-Enabled Research Cloud Computing Fellowship to MDR.

# 8 Conflicts of interest

None declared.

# References

1. J. E. Aagaard, J. H. Willis, and P. C. Phillips. Relaxed selection among duplicate floral regulatory genes in Lamiales. *Journal of Molecular Evolution*, 63(4):493–503, Oct. 2006.

2. C. Alonso-Blanco, J. Andrade, C. Becker, F. Bemm, J. Bergelson, K. M. Borgwardt, J. Cao, E. Chae, T. M. Dezwaan, W. Ding, J. R. Ecker, M. Exposito-Alonso, A. Farlow, J. Fitz, X. Gan, D. G. Grimm, A. M. Hancock, S. R. Henz, S. Holm, M. Horton, M. Jarsulic, R. A. Kerstetter, A. Korte, P. Korte, C. Lanz, C.-R. Lee, D. Meng, T. P. Michael, R. Mott, N. W. Muliyati, T. Nägele, M. Nagler, V. Nizhynska, M. Nordborg, P. Y. Novikova, F. X. Picó, A. Platzer, F. A. Rabanal, A. Rodriguez, B. A. Rowan, P. A. Salomé, K. J. Schmid, R. J. Schmitz, Seren, F. G. Sperone, M. Sudkamp, H. Svardal, M. M. Tanzer, D. Todd, S. L. Volchenboum, C. Wang, G. Wang, X. Wang, W. Weckwerth, D. Weigel, and X. Zhou. 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell*, 166(2):481–491, July 2016. Publisher: Elsevier.

3. D. Alvarez-Ponce and M. A. Fares. Evolutionary rate and duplicability in the *Arabidopsis thaliana* protein–protein interaction network. *Genome Biology and Evolution*, 4(12):1263–1274, Dec. 2012.

4. D. Alvarez-Ponce, F. Feyertag, and S. Chakraborty. Position matters: network centrality considerably impacts rates of protein evolution in the human protein–protein interaction network. *Genome Biology and Evolution*, 9(6):1742–1756, June 2017.

5. J. R. Auld, A. A. Agrawal, and R. A. Relyea. Re-evaluating the costs and limits of adaptive phenotypic plasticity. *Proceedings of the Royal Society B: Biological Sciences*, 277(1681):503–511, Feb. 2010. Publisher: Royal Society.

6. A. J. Betancourt and D. C. Presgraves. Linkage limits the power of natural selection in *Drosophila*. *Proceedings of the National Academy of Sciences*, 99(21):13616–13620, Oct. 2002. Publisher: Proceedings of the National Academy of Sciences.

7. S. J. Bush, P. X. Kover, and A. O. Urrutia. Lineage-specific sequence evolution and exon edge conservation partially explain the relationship between evolutionary rate and expression level in *A. thaliana*. *Molecular Ecology*, 24(12):3093–3106, 2015. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/mec.13221.

8. C. I. Castillo-Davis, S. L. Mekhedov, D. L. Hartl, E. V. Koonin, and F. A. Kondrashov. Selection for short introns in highly expressed genes. *Nature Genetics*, 31(4):415–418, Aug. 2002. Number: 4 Publisher: Nature Publishing Group.

9. B. Charlesworth. The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genetical Research*, 63(3):213–227, June 1994.

10. S. Chen, Y. Zhou, Y. Chen, and J. Gu. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34(17):i884–i890, Sept. 2018.

11. C.-Y. Cheng, V. Krishnakumar, A. P. Chan, F. Thibaud-Nissen, S. Schobel, and C. D. Town. Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *The Plant Journal*, 89(4):789–804, 2017. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/tpj.13415.

12. P. Cingolani, V. M. Patel, M. Coon, T. Nguyen, S. J. Land, D. M. Ruden, and X. Lu. Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift. *Frontiers in Genetics*, 3, 2012. Publisher: Frontiers Media SA.

13. P. Cingolani, A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D. M. Ruden. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly*, 6(2):80–92, Apr. 2012.

14. M. Colombo, H. Laayouni, B. M. Invergo, J. Bertranpetit, and L. Montanucci. Metabolic flux is a determinant of the evolutionary rates of enzyme-encoding genes. *Evolution*, 68(2):605–613, 2014. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/evo.12262.

15. L. Cooper, A. Meier, M.-A. Laporte, J. L. Elser, C. Mungall, B. T. Sinn, D. Cavaliere, S. Carbon, N. A. Dunn, B. Smith, B. Qu, J. Preece, E. Zhang, S. Todorovic, G. Gkoutos, J. H. Doonan, D. W. Stevenson, E. Arnaud, and P. Jaiswal. The Planteome database: an integrated resource for reference ontologies, plant genomics and phenomics. *Nucleic Acids Research*, 46(D1):D1168–D1180, Jan. 2018.

16. F. Crameri. Geodynamic diagnostics, scientific visualisation and StagLab 3.0. *Geoscientific Model Development*, 11(6):2541–2562, June 2018. Publisher: Copernicus GmbH.

17. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, and . G. P. A. Group. The variant call format and VCFtools. *Bioinformatics*, 27(15):2156–2158, Aug. 2011. Publisher: Oxford Academic.

18. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, and H. Li. Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2):giab008, Feb. 2021.

19. J. C. Davis and D. A. Petrov. Preferential duplication of conserved proteins in eukaryotic genomes. *PLOS Biology*, 2(3):e55, Mar. 2004. Publisher: Public Library of Science.

20. T. J. DeWitt, A. Sih, and D. S. Wilson. Costs and limits of phenotypic plasticity. *Trends in Ecology & Evolution*, 13(2):77–81, Feb. 1998.

21. D. A. Drummond, A. Raval, and C. O. Wilke. A single determinant dominates the rate of yeast protein evolution. *Molecular Biology and Evolution*, 23(2):327–337, Feb. 2006.

22. L. Duret and D. Mouchiroud. Determinants of substitution rates in mammalian genes: expression pattern affects selection intensity but not mutation rate. *Molecular Biology and Evolution*, 17(1):68–070, Jan. 2000.

23. E. Eisenberg and E. Y. Levanon. Human housekeeping genes are compact. *Trends in Genetics*, 19(7):362–365, July 2003.

24. D. M. Emms and S. Kelly. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, 20(1):238, Nov. 2019.

25. P. Ewels, M. Magnusson, S. Lundin, and M. Käller. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32(19):3047–3048, Oct. 2016.

26. K. Fukushima and D. D. Pollock. Amalgamated cross-species transcriptomes reveal organ-specific propensity in gene expression evolution. *Nature Communications*, 11(1):4459, Sept. 2020. Number: 1 Publisher: Nature Publishing Group.

27. B. Gaut, L. Yang, S. Takuno, and L. E. Eguiarte. The patterns and causes of variation in plant nucleotide substitution rates. *Annual Review of Ecology, Evolution, and Systematics*, 42(1):245–266, Dec. 2011.

28. S. Glémin. Mating systems and the efficacy of selection at the molecular level. *Genetics*, 177(2):905–916, Oct. 2007.

29. S. C. Groen, I. Ćalić, Z. Joly-Lopez, A. E. Platts, J. Y. Choi, M. Natividad, K. Dorph, W. M. Mauck, B. Bracken, C. L. U. Cabral, A. Kumar, R. O. Torres, R. Satija, G. Vergara, A. Henry, S. J. Franks, and M. D. Purugganan. The strength and pattern of natural selection on gene expression in rice. *Nature*, 578(7796):572–576, Feb. 2020. Number: 7796 Publisher: Nature Publishing Group.

30. Y.-F. Huang. Dissecting genomic determinants of positive selection with an evolution-guided regression model. *Molecular Biology and Evolution*, 39(1):msab291, Jan. 2022.

31. C. D. Huber, B. Y. Kim, C. D. Marsden, and K. E. Lohmueller. Determining the factors driving selective effects of new nonsynonymous mutations. *Proceedings of the National Academy of Sciences*, 114(17):4465–4470, Apr. 2017. Publisher: Proceedings of the National Academy of Sciences.

32. B. G. Hunt, L. Ometto, Y. Wurm, D. Shoemaker, S. V. Yi, L. Keller, and M. A. D. Goodisman. Relaxed selection is a precursor to the evolution of phenotypic plasticity. *Proceedings of the National Academy of Sciences*, 108(38):15936–15941, Sept. 2011. Publisher: National Academy of Sciences Section: Biological Sciences.

33. I. K. Jordan, Y. I. Wolf, and E. V. Koonin. Duplicated genes evolve slower than singletons despite the initial rate increase. *BMC Evolutionary Biology*, 4(1):22, July 2004.

34. E. B. Josephs, S. I. Wright, J. R. Stinchcombe, and D. J. Schoen. The relationship between selection, network connectivity, and regulatory variation within a population of *Capsella grandiflora*. *Genome Biology and Evolution*, 9(4):1099–1109, Apr. 2017.

35. K. Katoh and D. M. Standley. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4):772–780, Apr. 2013. Publisher: Oxford Academic.

36. T. J. Kawecki. Accumulation of Deleterious Mutations and the Evolutionary Cost of Being a Generalist. *The American Naturalist*, 144(5):833–838, Nov. 1994. Publisher: The University of Chicago Press.

37. S. Kim. ppcor: an R package for a fast calculation to semi-partial correlation coefficients. *Communications for statistical applications and methods*, 22(6):665–674, Nov. 2015.

38. E. V. Koonin. Are there laws of genome evolution? *PLOS Computational Biology*, 7(8):e1002173, Aug. 2011. Publisher: Public Library of Science.

39. K. L. Korunes and K. Samuk. pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Molecular Ecology Resources*, 21(4):1359–1368, 2021. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1755-0998.13326.

40. N. Kryuchkova-Mostacci and M. Robinson-Rechavi. A benchmark of gene expression tissue-specificity metrics. *Briefings in Bioinformatics*, 18(2):205–214, Mar. 2017.

41. R. Lanfear, S. Y. W. Ho, T. Jonathan Davies, A. T. Moles, L. Aarssen, N. G. Swenson, L. Warman, A. E. Zanne, and A. P. Allen. Taller plants have lower rates of molecular evolution. *Nature Communications*, 4(1):1879, May 2013. Number: 1 Publisher: Nature Publishing Group.

42. C. H. Langley, K. Stevens, C. Cardeno, Y. C. G. Lee, D. R. Schrider, J. E. Pool, S. A. Langley, C. Suarez, R. B. Corbett-Detig, B. Kolaczkowski, S. Fang, P. M. Nista, A. K. Holloway, A. D. Kern, C. N. Dewey, Y. S. Song, M. W. Hahn, and D. J. Begun. Genomic variation in natural populations of *Drosophila melanogaster*. *Genetics*, 192(2):533–598, Oct. 2012.

43. A. M. Larracuente, T. B. Sackton, A. J. Greenberg, A. Wong, N. D. Singh, D. Sturgill, Y. Zhang, B. Oliver, and A. G. Clark. Evolution of protein-coding genes in *Drosophila*. *Trends in Genetics*, 24(3):114–123, Mar. 2008.

44. J. T. Leek, R. B. Scharpf, H. C. Bravo, D. Simcha, B. Langmead, W. E. Johnson, D. Geman, K. Baggerly, and R. A. Irizarry. Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature Reviews Genetics*, 11(10):733–739, Oct. 2010. Number: 10 Publisher: Nature Publishing Group.

45. J. T. Leek and J. D. Storey. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLOS Genetics*, 3(9):e161, Sept. 2007. Publisher: Public Library of Science.

46. H. Li and R. Durbin. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14):1754–1760, July 2009.

47. Y.-S. Lin, W.-L. Hsu, J.-K. Hwang, and W.-H. Li. Proportion of solvent-exposed amino acids in a protein and rate of protein evolution. *Molecular Biology and Evolution*, 24(4):1005–1011, Apr. 2007.

48. M. Lynch and J. S. Conery. The evolutionary fate and consequences of duplicate genes. *Science*, 290(5494):1151–1155, Nov. 2000. Publisher: American Association for the Advancement of Science.

49. R. R. Masalia, A. J. Bewick, and J. M. Burke. Connectivity in gene coexpression networks negatively correlates with rates of molecular evolution in flowering plants. *PLOS ONE*, 12(7):e0182289, July 2017. Publisher: Public Library of Science.

50. K. McGuigan, J. M. Collet, S. L. Allen, S. F. Chenoweth, and M. W. Blows. Pleiotropic mutations are subject to strong stabilizing selection. *Genetics*, 197(3):1051–1062, July 2014.

51. M. Mondragón-Palomino, B. C. Meyers, R. W. Michelmore, and B. S. Gaut. Patterns of positive selection in the complete NBS-LRR gene family of *Arabidopsis thaliana*. *Genome Research*, 12(9):1305–1315, Sept. 2002. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

52. A. F. Moutinho, A. Eyre-Walker, and J. Y. Dutheil. Strong evidence for the adaptive walk model of gene evolution in *Drosophila* and *Arabidopsis*. *PLOS Biology*, 20(9):e3001775, Sept. 2022. Publisher: Public Library of Science.

53. D. Mukherjee, A. Mukherjee, and T. C. Ghosh. Evolutionary rate heterogeneity of primary and secondary metabolic pathway genes in *Arabidopsis thaliana*. *Genome Biology and Evolution*, 8(1):17–28, Jan. 2016.

54. C. J. Murren, J. R. Auld, H. Callahan, C. K. Ghalambor, C. A. Handelsman, M. A. Heskel, J. G. Kingsolver, H. J. Maclean, J. Masel, H. Maughan, D. W. Pfennig, R. A. Relyea, S. Seiter, E. Snell-Rood, U. K. Steiner, and C. D. Schlichting. Constraints on the evolution of phenotypic plasticity: limits and costs of phenotype and plasticity. *Heredity*, 115(4):293–301, Oct. 2015. Bandiera_abtest: a Cc_license_type: cc_y Cg_type: Nature Research Journals Number: 4 Primary_atype: Reviews Publisher: Nature Publishing Group Subject_term: Evolutionary theory Subject_term_id: evolutionary-theory.

55. N. Mähler, J. Wang, B. K. Terebieniec, P. K. Ingvarsson, N. R. Street, and T. R. Hvidsten. Gene co-expression network connectivity is an important determinant of selective constraint. *PLoS Genetics*, 13(4):e1006402, Apr. 2017.

56. M. Nei and T. Gojobori. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, 3(5):418–426, Sept. 1986.

57. V. Nembaware, K. Crum, J. Kelso, and C. Seoighe. Impact of the presence of paralogs on sequence divergence in a set of mouse-human orthologs. *Genome Research*, 12(9):1370–1376, Sept. 2002. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

58. M. Nordborg, T. T. Hu, Y. Ishino, J. Jhaveri, C. Toomajian, H. Zheng, E. Bakker, P. Calabrese, J. Gladstone, R. Goyal, M. Jakobsson, S. Kim, Y. Morozov, B. Padhukasahasram, V. Plagnol, N. A. Rosenberg, C. Shah, J. D. Wall, J. Wang, K. Zhao, T. Kalbfleisch, V. Schulz, M. Kreitman, and J. Bergelson. The pattern of polymorphism in *Arabidopsis thaliana*. *PLOS Biology*, 3(7):e196, May 2005. Publisher: Public Library of Science.

59. S. Ossowski, K. Schneeberger, J. I. Lucas-Lledó, N. Warthmann, R. M. Clark, R. G. Shaw, D. Weigel, and M. Lynch. The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science*, 327(5961):92–94, Jan. 2010. Publisher: American Association for the Advancement of Science.

60. T. Paape, T. Bataillon, P. Zhou, T. J. Y. Kono, R. Briskine, N. D. Young, and P. Tiffin. Selection, genome-wide fitness effects and evolutionary rates in the model legume *Medicago truncatula*. *Molecular Ecology*, 22(13):3525–3538, 2013. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/mec.12329.

61. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, and C. Kingsford. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, 14(4):417–419, Apr. 2017. Number: 4 Publisher: Nature Publishing Group.

62. B. L. Payne and D. Alvarez-Ponce. Higher rates of protein evolution in the self-fertilizing plant *Arabidopsis thaliana* than in the out-crossers *Arabidopsis lyrata* and *Arabidopsis halleri*. *Genome Biology and Evolution*, 10(3):895–900, Mar. 2018.

63. T. L. Pedersen and F. Crameri. *scico: Colour Palettes Based on the Scientific Colour-Maps*. R, 2022.

64. M. Pigliucci and A. Kolodynska. Phenotypic plasticity to light intensity in *Arabidopsis thaliana*: invariance of reaction norms and phenotypic integration. *Evolutionary Ecology*, 16(1):27–47, Jan. 2002.

65. J. B. Plotkin and H. B. Fraser. Assessing the determinants of evolutionary rates in the presence of noise. *Molecular Biology and Evolution*, 24(5):1113–1121, May 2007.

66. E. P. C. Rocha. The quest for the universals of protein evolution. *Trends in Genetics*, 22(8):412–416, Aug. 2006.

67. S. M. Scheiner. Genetics and evolution of phenotypic plasticity. *Annual Review of Ecology and Systematics*, 24(1):35–68, 1993. _eprint: https://doi.org/10.1146/annurev.es.24.110193.000343.

68. C. D. Schlichting and H. Smith. Phenotypic plasticity: linking molecular mechanisms with evolutionary outcomes. *Evolutionary Ecology*, 16(3):189–211, May 2002.

69. H. Schneider. Characterization, costs, cues, and future perspectives of phenotypic plasticity. *Annals of botany*, June 2022.

70. T. Slotte, T. Bataillon, T. T. Hansen, K. St. Onge, S. I. Wright, and M. H. Schierup. Genomic determinants of protein evolution and polymorphism in *Arabidopsis*. *Genome Biology and Evolution*, 3:1210–1219, Jan. 2011.

71. E. C. Snell-Rood, J. D. Van Dyken, T. Cruickshank, M. J. Wade, and A. P. Moczek. Toward a population genetic framework of developmental evolution: the costs, limits, and consequences of phenotypic plasticity. *BioEssays*, 32(1):71–81, 2010. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/bies.200900132.

72. N. Stoletzki and A. Eyre-Walker. Estimation of the neutrality index. *Molecular Biology and Evolution*, 28(1):63–70, Jan. 2011.

73. M. Suyama, D. Torrents, and P. Bork. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, 34(suppl_2):W609–W612, July 2006.

74. F. Tajima. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123(3):585–595, Nov. 1989.

75. S. Takuno and B. S. Gaut. Body-methylated genes in *Arabidopsis thaliana* are functionally important and evolve slowly. *Molecular Biology and Evolution*, 29(1):219–227, Jan. 2012.

76. A. O. Urrutia and L. D. Hurst. Codon usage bias covaries with expression breadth and the rate of synonymous evolution in humans, but this is not evidence for selection. *Genetics*, 159(3):1191–1199, Nov. 2001.

77. A. O. Urrutia and L. D. Hurst. The signature of selection mediated by expression on human genes. *Genome Research*, 13(10):2260–2264, Oct. 2003. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

78. J. Van Buskirk and U. K. Steiner. The fitness costs of developmental canalization and plasticity. *Journal of Evolutionary Biology*, 22(4):852–860, 2009. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1420-9101.2009.01685.x.

79. J. D. Van Dyken and M. J. Wade. The genetic signature of conditional expression. *Genetics*, 184(2):557–570, Feb. 2010.

80. M. Van Kleunen and M. Fischer. Constraints on the evolution of adaptive phenotypic plasticity in plants. *New Phytologist*, 166(1):49–60, 2005. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1469-8137.2004.01296.x.

81. H.-c. Wang, G. A. C. Singer, and D. A. Hickey. Mutational bias affects protein evolution in flowering plants. *Molecular Biology and Evolution*, 21(1):90–96, Jan. 2004.

82. L. C. Wheeler, J. F. Walker, J. Ng, R. Deanna, A. Dunbar-Wallis, A. Backes, P. H. Pezzi, M. V. Palchetti, H. M. Robertson, A. Monaghan, L. B. de Freitas, G. E. Barboza, E. Moyroud, and S. D. Smith. Transcription factors evolve faster than their structural gene targets in the flavonoid pigment pathway. *Molecular Biology and Evolution*, 39(3):msac044, Mar. 2022.

83. M. C. Whitlock. The Red Queen Beats the Jack-Of-All-Trades: The Limitations on the Evolution of Phenotypic Plasticity and Niche Breadth. *The American Naturalist*, 148:S65–S77, 1996. Publisher: [University of Chicago Press, American Society of Naturalists].

84. E. E. Winter, L. Goodstadt, and C. P. Ponting. Elevated rates of protein secretion, evolution, and disease among tissue-specific genes. *Genome Research*, 14(1):54–61, Jan. 2004. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

85. S. I. Wright, B. Lauga, and D. Charlesworth. Rates and patterns of molecular evolution in inbred and outbred *Arabidopsis*. *Molecular Biology and Evolution*, 19(9):1407–1420, Sept. 2002.

86. S. I. Wright, C. B. K. Yau, M. Looseley, and B. C. Meyers. Effects of gene expression on molecular evolution in *Arabidopsis thaliana* and *Arabidopsis lyrata*. *Molecular Biology and Evolution*, 21(9):1719–1726, Sept. 2004.

87. Z. Wu, X. Cai, X. Zhang, Y. Liu, G.-b. Tian, J.-R. Yang, and X. Chen. Expression level is a major modifier of the fitness landscape of a protein coding gene. *Nature Ecology & Evolution*, 6(1):103–115, Jan. 2022. Number: 1 Publisher: Nature Publishing Group.

88. I. Yanai, H. Benjamin, M. Shmoish, V. Chalifa-Caspi, M. Shklar, R. Ophir, A. Bar-Even, S. Horn-Saban, M. Safran, E. Domany, D. Lancet, and O. Shmueli. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics*, 21(5):650–659, Mar. 2005.

89. J. Yang, Z. Gu, and W.-H. Li. Rate of protein evolution versus fitness effect of gene deletion. *Molecular Biology and Evolution*, 20(5):772–774, May 2003.

90. L. Yang and B. S. Gaut. Factors that contribute to variation in evolutionary rate among *Arabidopsis* genes. *Molecular Biology and Evolution*, 28(8):2359–2369, Aug. 2011.

91. Z. Yang. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution*, 24(8):1586–1591, Aug. 2007.

92. H. Zhang, F. Zhang, Y. Yu, L. Feng, J. Jia, B. Liu, B. Li, H. Guo, and J. Zhai. A comprehensive online database for exploring 20,000 public *Arabidopsis* RNA-seq libraries. *Molecular Plant*, 13(9):1231–1233, Sept. 2020.

93. J. Zhang and J.-R. Yang. Determinants of the rate of protein sequence evolution. *Nature Reviews Genetics*, 16(7):409–420, July 2015. Number: 7 Publisher: Nature Publishing Group.

94. L. Zhang and W.-H. Li. Mammalian housekeeping genes evolve more slowly than tissue-specific genes. *Molecular Biology and Evolution*, 21(2):236–239, Feb. 2004.

95. L. Zhang, T. J. Vision, and B. S. Gaut. Patterns of nucleotide substitution among simultaneously duplicated gene pairs in *Arabidopsis thaliana*. *Molecular Biology and Evolution*, 19(9):1464–1473, Sept. 2002.

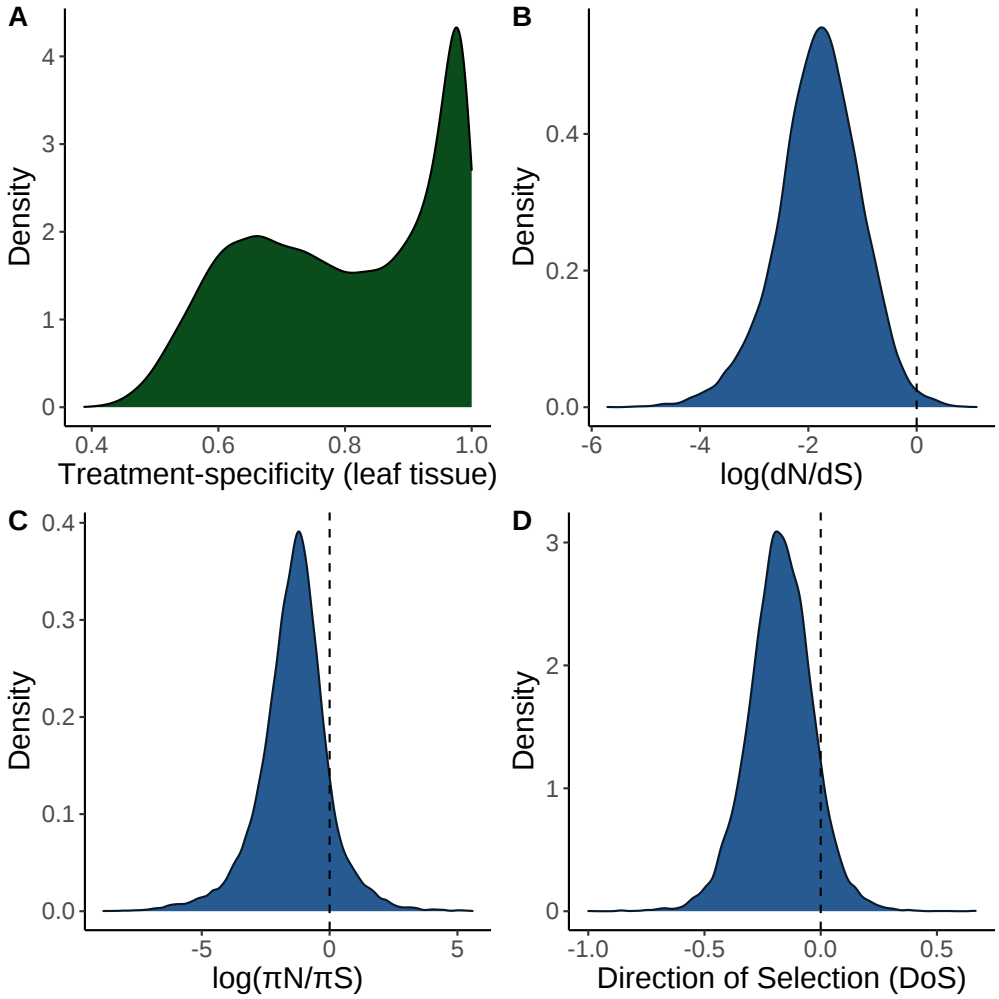


**Figure 2.** Density plots of key variables measured in this study. (**A**) Distribution of treatment specificity in leaf tissue expression across all genes included in this study. The area underneath the curve in a given interval of treatment specificity represents the proportion of genes in this study 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 $\pi_N/\pi_S$ 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 propotion of genes with DoS ¿ 0, which is interpreted as evidence of adaptive evolution.

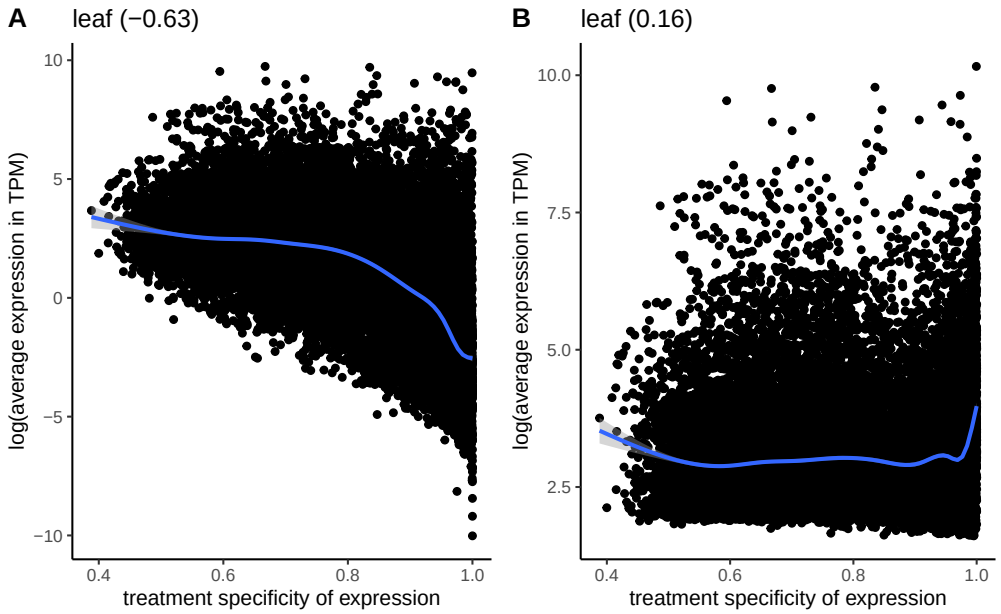


**Figure 3.** Correlation between the average expression in transcripts per million (TPM) and treatment specificity of genes when samples with low expression (< 5 TPM) are included (**A**) vs excluded (**B**). Expression level and treatment specificity were calculated using only data from leaf tissue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

**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 | $7.6 \times 10^{-31}$ |
| Post | $dN$ | 0.13 | $6.9 \times 10^{-50}$ |
| Pre | $\pi_N$ | 0.10 | $1.2 \times 10^{-62}$ |
| Post | $\pi_N$ | 0.16 | $3.9 \times 10^{-128}$ |
| Pre | Tajima's D | 0.03 | $3.1 \times 10^{-7}$ |
| Post | Tajima's D | 0.04 | $6.6 \times 10^{-10}$ |
| Pre | DoS | 0.04 | $2.3 \times 10^{-6}$ |
| Post | DoS | 0.05 | $2.0 \times 10^{-8}$ |

[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.

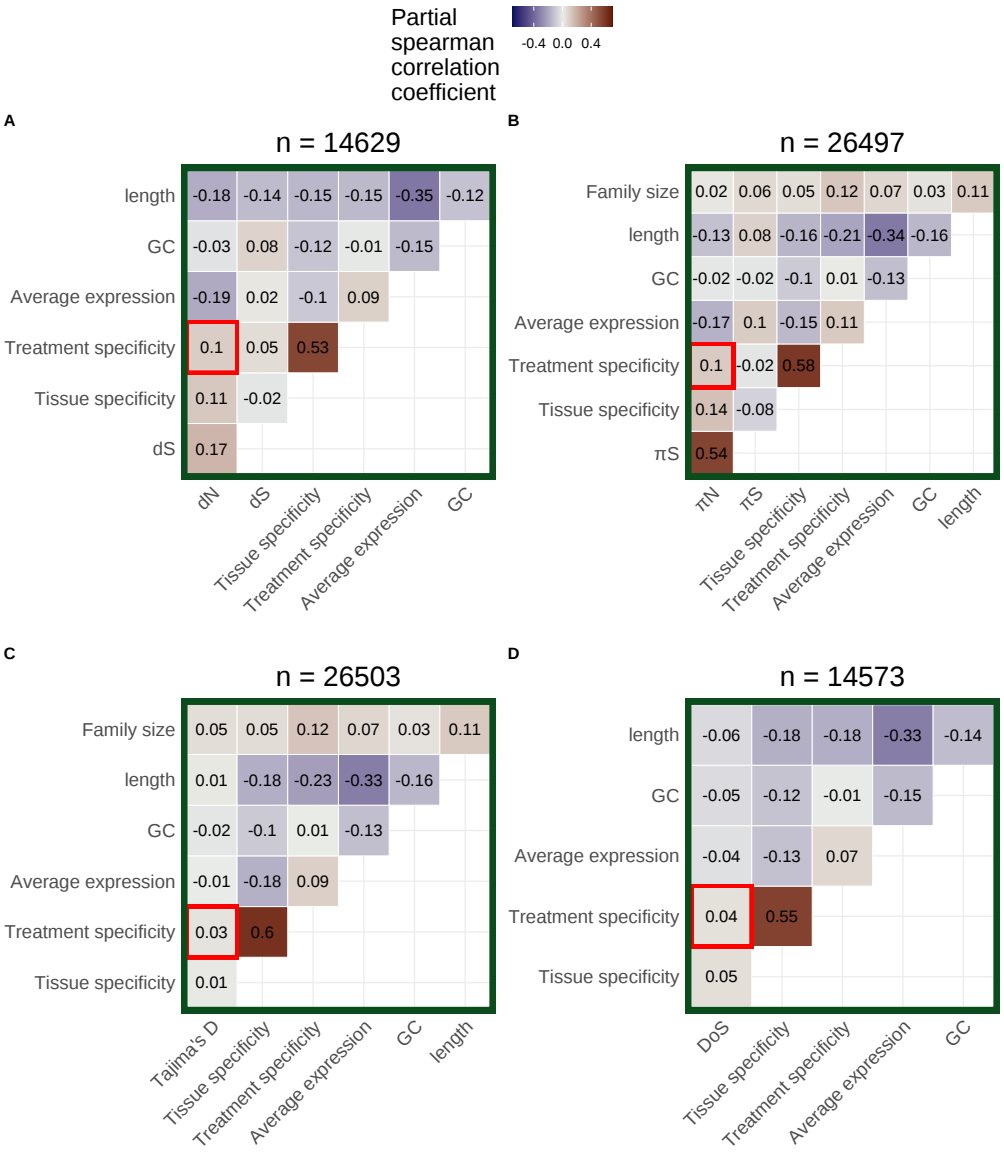


**Figure 4.** Partial correlation analysis including either (**A**) $dN$, (**B**) $\pi_N$, (**C**) Tajima's D, or (**D**) direction of selection (DoS) as a covariate. Average expression excludes values ¡ 5 TPM and was calculated using only leaf tissue samples. Treatment specificity was also calculated 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.

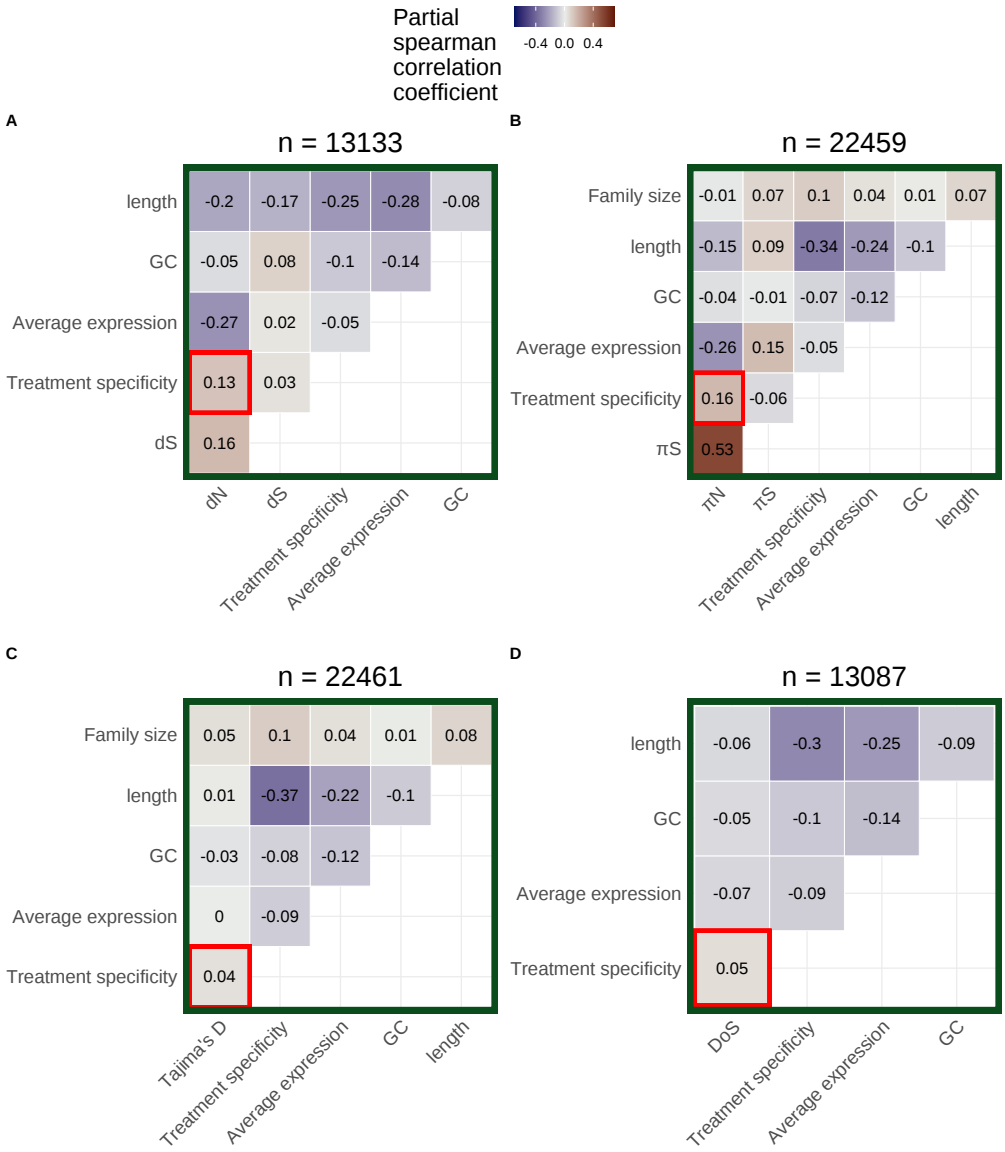


**Figure 5.** Partial correlations for (**A**) $dN$, (**B**) $\pi_N$, (**C**) Tajima's D, and (**D**) direction of selection (DoS) based on leaf tissue data subset after applying SVA. Data was further subset to include only treatment groups with data from more than one study before applying SVA. Average 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.