Abundant *trans*-driven variation in detoxification gene expression in the extreme generalist herbivore *Tetranychus urticae*

Andre H Kurlovs[1,2¶], Berdien De Beer[1¶], Meiyuan Ji[2¶], Marilou Vandenhole[1], Tim De Meyer[3], René Feyereisen[1,4], Richard M Clark[2,5*], Thomas Van Leeuwen[1*]

[1] Department of Plants and Crops, Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium

[2] School of Biological Sciences, University of Utah, Salt Lake City, United States of America

[3] Department of Data Analysis and Mathematical Modelling, Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium

[4] Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej, Copenhagen, Denmark

[5] Henry Eyring Center for Cell and Genome Science, University of Utah, Salt Lake City, United States of America

*Corresponding authors

E-mail: clark@biology.utah.edu (RMC); Thomas.VanLeeuwen@UGent.be (TVL)

¶ These authors contributed equally to this work.

Short title: Gene expression regulation in *Tetranychus urticae*

Abstract

The extreme adaptation potential of the generalist herbivore *Tetranychus urticae* (the two-spotted spider mite) to pesticides as well as diverse host plants has been associated with clade-specific gene expansions in known detoxifying enzyme families, and with extensive and rapid transcriptional responses. However, how this broad transcriptional potential is regulated remains largely unknown. Using a parental/F1 design in which four inbred strains were crossed to a common inbred strain, we assessed the genetic basis and inheritance of gene expression variation in *T. urticae*. Mirroring known phenotypic variation in the progenitor strains of the inbreds, we confirmed that the inbred strains we created were genetically distinct, varied markedly in pesticide resistance, and also captured variation in host plant fitness commonly observed in this species. By examining differences in gene expression between parents and allele-specific expression in F1s, we found that variation in RNA abundance was more often explained in *trans* as compared to *cis*, with the former associated with dominance in inheritance. Strikingly, in a gene ontology analysis, detoxification genes of the cytochrome P450 monooxygenase (CYP) family, as well as dioxygenases (DOGs) acquired from horizontal gene transfer from fungi, were specifically enriched at the extremes of *trans*-driven up- and downregulation. In particular, a clade of CYPs with documented broad substrate-specificity including multiple pesticides, as well as DOGs that have recently been shown to have broad substrate specificity against plant specialized compounds, were exceptionally highly upregulated as a result of *trans* effects (or in some cases synergism of *cis* and *trans*) in the most multi-pesticide resistant strains. Collectively, our findings highlight the potential importance of *trans*-driven expression variation in genes associated with xenobiotic metabolism and host plant use for rapid adaptation in *T. urticae*, and also suggest modular control of these genes, a regulatory architecture that might ameliorate negative pleiotropic effects.

Author summary

The two-spotted spider mite, *Tetranychus urticae,* is a generalist herbivore and pest of diverse crops globally. In response to the plethora of chemicals used for its control, the species rapidly evolves pesticide resistance. Further, experimental evolution studies with *T. urticae* populations have demonstrated adaptation to challenging host plants in as few as five generations. The adaptation of *T. urticae* to pesticides and host plants has been associated with large transcriptome changes, including for genes associated with detoxification of pesticides and toxic plant compounds. Nevertheless, the basis of the observed transcriptome variation has remained largely unknown. Here, we examined the genetic control and inheritance of expression differences among five inbred *T. urticae* strains, including several with histories of intense pesticide selection. With a parental/F1 experimental design, we found that *trans* effects were common in explaining variation in detoxification gene expression, with the *trans*-driven upregulation of a subset of cytochrome P450 monooxygenases of broad substrate specificity especially striking in the most pesticide resistant strains. Our findings suggest that genetic variation acting with dominant or additive inheritance to impact the regulation of modules of detoxification genes may be an important target of selection during rapid pesticide and host plant evolution in herbivores.

Introduction

Genetic variation as a requisite for evolutionary change through selection, recombination and drift lies at the basis of our current understanding of evolution [1,2]. Nucleotide changes in coding sequences were among the first to be associated with phenotypic differences of adaptive significance. Subsequently, advances in genomic methods have facilitated a broader understanding of how other types of coding sequence variation – such as structural or copy number variation – impact phenotypes [3,4]. Likewise, a number of specific instances of regulatory variation in genic regions impacting gene expression phenotypes of adaptive relevance have been characterized in detail [5–8]. More recently, advances in experimental designs and methods have allowed the genetic control of gene expression variation to be assessed genome-wide [9,10], offering the possibility of a more comprehensive understanding of how such variation arises and is selected upon to impact phenotypes [11].

One widely used approach to understand the genetic control of gene expression variation, as well as its associated inheritance, involves comparing expression between parents and F1 offspring [9,12–21]. If a difference in gene expression between two individuals or parental strains is due to *cis* effects only (i.e., changes in promoter or enhancer regions), the same difference in expression will be observed between the respective parental allele copies in F1 offspring (i.e., allele-specific expression, or ASE). However, if expression differences between parents are not observed in F1 ASE (or parental and ASE patterns contrast), *trans* effects contribute to regulatory variation [9]. This type of parental/F1 design also enables the degree of dominance – an important factor in understanding responses to selection – to be inferred [14,19–21]. A number of studies in diploid organisms, often using genetic models selected in part because of their extensive genomic resources, have revealed trends about the frequency of *cis* and *trans* effects, their interactions, as well as associations with inheritance type. For example, in many instances, expression variation more often results from *trans* versus *cis* regulation in intraspecific comparisons, with the opposite pattern observed for interspecific ones ([22], and references therein). Further, while *trans*

regulation is associated with dominance or non-additivity in gene expression inheritance, genes with *cis* effects are enriched for additive inheritance [14,16,19–21].

However, despite the resolution and comprehensive nature of parental/F1 and related experimental designs, connecting the genetic variation that impacts gene expression to adaptation has remained challenging. In part, this results from a lack of understanding of underlying selective forces. A classic example of rapid adaptation for which the agent(s) are known explicitly is pesticide resistance evolution. This phenotype has been studied extensively in arthropods, including in herbivores [23–26]. For the herbivorous insects and mites that consume plant tissue, chemical control by pesticides is widely used. Resistance evolution is a major obstacle for agriculture, and understanding inheritance (dominance or lack thereof) can be critical to inform approaches for resistance management [27–29]. Further, albeit with less precision in space, time, or nature of the selective agent(s), adaptation of herbivores to their host plants has also provided well-known examples of evolution in action [30–34]. Finally, a number of gene families have been associated with the detoxification, sequestration, or transport and excretion of pesticides, and are either known or are likely to be critical in plant host adaptation as well (i.e., to overcome toxic plant-produced specialized compounds, although plants deploy other defense strategies as well) [33–41]. Therefore, many selective agents are known for adaptation in herbivores, sometimes exactly (pesticides), as are genes and gene families that are strong candidates to underlie rapid phenotypic evolution.

In this study, we have determined the genetic basis and inheritance of gene expression variation among strains of the two-spotted spider mite, *Tetranychus urticae*. This generalist (polyphagous) mite is at the extreme end of the generalist-to-specialist spectrum, as it has been documented to feed on a staggering 1,100 plant species from more than 250 plant families, including more than 150 crops [42]. In addition, *T. urticae* is well known for developing acaricide resistance in as fast as a few years after the introduction of most acaricides used for its control (acaricides are pesticides that are active against mites, which belong to the Acari within Arthropoda subphylum Chelicerata) [25]. The extreme adaptation potential of this species has been associated with clade-specific gene expansions in known detoxifying

enzyme families, such as cytochrome P450 monooxygenases (CYPs), mu-class glutathione-S-transferases (GSTs), a new clade of carboxyl-choline esterases (CCEs), and an unexpectedly large repertoire of ATP-binding cassette (ABC) and major facilitator superfamily (MFS) transporters [33,36,42]. In addition, it acquired several putative detoxification genes by horizontal gene transfer that subsequently proliferated in the genome [43]. These include intradiol-ring cleavage dioxygenases (DOGs) that have recently been shown to cleave an unexpectedly wide range of aromatic plant defense specialized compounds [41], as well as bacterial UDP-glucuronosyltransferases (UGTs) [44], demonstrating the adaptive advantage of horizontal gene transfer in the context of detoxification and host plant use [42,43,45]. Gene-expression studies have revealed large and similar transcriptomic differences in these gene families between acaricide susceptible and resistant *T. urticae* strains, as well as after adaptation to new host plants [33,34,46], suggesting a link between host plant use and acaricide resistance. The extensive and rapid transcriptional reprogramming associated with the exceptional adaptation potential of *T. urticae* raises several questions about the mechanisms of gene regulation in polyphagous herbivores. Given the often very large population sizes of this species in agricultural settings, selection acting on *cis* variants at many loci is plausible. However, the often coordinated and rapid responses in gene expression patterns observed between strains after selection by both acaricides and host plants [33,34] raises the alternative possibility of selection acting on *trans*-acting factors that serve as master regulators of detoxification/adaptation pathways.

As opposed to many herbivores, *T. urticae* has a short life cycle and a high-quality genome, and is easy to inbreed [47]. Therefore, it is optimally suited for a study relying on F1s from parental strain crosses to determine *cis*/*trans* contributions to and inheritance of gene expression. Here, we selected four genetically divergent *T. urticae* strains with different evolutionary histories, including two highly multi-acaricide resistant strains, and inbred them. We then crossed the inbred strains to a common inbred parental strain that was comparatively susceptible to most acaricides tested. Using parental/F1 experimental designs, we observed extensive expression differences among strains. We found that gene expression differences arising from *trans* variation were common, including for subsets of genes in

multiple detoxification gene families, indicative of modular control of detoxification and plant host associated pathways in this generalist mite pest. Notably, *trans* control contributed prominently to the exceptionally high expression of a clade of CYPs with broad substrate specificity in the acaricide resistant strains.

Results

Five inbred strains harbor extensive genetic variation

Strains of *T. urticae* used in this study were selected based on their different life histories and acaricide resistance status (see Methods). Briefly, strains MR-VP and MAR-AB originated from European strains previously found to be multi-acaricide resistant [33,48], strain JP-RR originated from a Japanese strain resistant to several classes of mitochondrial electron transport inhibitors (METI) [49], while strain ROS-IT was recently collected in Italy and was highly resistant to abamectin [50]. Finally, strain SOL-BE originated from a Belgian potato field which had not undergone any acaricide treatment. To fix intra-population genetic variation, a prerequisite for our downstream analyses, each of the five strains was inbred by mother-son crossings for at least seven generations. The subsequent inbred strains, denoted MR-VPi, MAR-ABi, JP-RRi, ROS-ITi, and SOL-BEi, were used for all subsequent analyses, including genome sequencing. Resulting Illumina DNA-seq reads for each strain were aligned to the ~90 Mb *T. urticae* three-chromosome reference assembly [51], and median coverages ranged from 98- (MAR-ABi) to 430-fold (ROS-ITi). Among the five strains, variant prediction revealed 1.19 million high-quality variants (SNPs and small indels), consistent with prior studies that have revealed substantial variation among *T. urticae* strains [51–54]. Further, more than 98% of SNP sites were fixed in each of the five strains, confirming earlier reports that ~7 generations of inbreeding in this haplodiploid species results in isogenic (or nearly isogenic) strains [52]. To assess genetic relationships among strains, we performed a principal component analysis (PCA) using the SNP data (Fig 1A). While the inbreeding of the five strains was performed as part of the current study, the SOL-BEi and MAR-ABi strains, and respective genomic sequence

and parental RNA-seq data generated for these strains, was used by De Beer et al. [55] in a bulk segregant analysis (BSA) study that identified two quantitative trait loci (QTL) for resistance to the acaricide fenbutatin oxide [55]. Our PCA using all strains confirmed De Beer et al.'s [55] observation that SOL-BEi and MAR-ABi are genetically distinct, and moreover we found that SOL-BEi was highly divergent from each of the other three inbred strains. Two strains, MR-VPi and JP-RRi, did cluster nearby each other in the PCA analysis, but were nevertheless distinct along both PC axes 1 and 2, which explained 29.9% and 25.1% of the variance in this analysis, respectively.

The inbred strains have diverse and contrasting resistance profiles

In *T. urticae*, high levels of acaricide resistance and histories of exposure are often associated with variation in expression of genes encoding detoxification enzymes and transporters. To assess whether the inbred strains vary in acaricide resistance and life histories, we evaluated their levels of resistance to 12 acaricides belonging to multiple classes (Fig 2A and S1 Table), and also determined their distribution of target-site resistance mutations (Fig 2A and S2 Table). The five strains were distinguishable in their resistance to most acaricide classes, although fewer differences were observed for the METI-II and METI-III compounds. Across classes, several trends stood out. MR-VPi and MAR-ABi, for which progenitor populations were previously reported to be multi-acaricide resistant, exhibited the highest levels of resistance to compounds in multiple acaricide classes (Fig 2A). This was especially true for Na^+ channel modulators, GluCl allosteric modulators, inhibitors of mitochondrial ATP synthase, and METI-I compounds. Compared to MR-VPi, which is notably more resistant to METI-Is, MAR-ABi exhibited higher resistance to other classes, especially for inhibitors of mitochondrial ATP synthase (this study and De Beer et al. [55]). In contrast to MR-VPi and MAR-ABi, JP-RRi showed intermediate resistance to most acaricide classes, with modestly higher resistance levels to METI-II compounds. While the resistance profile of SOL-BEi was similar to JP-RRi, it was more resistant to the Na^+ channel modulator

bifenthrin. Unexpectedly, although ROS-ITi showed high susceptibility to most compounds tested, its level of resistance to dicofol (unknown target) was the highest among all strains (Fig 2A).

To further test for life histories that include exposure to acaricides, we used the genomic sequencing data and variant calls (Fig 1A) to survey target-site resistance mutations across strains (Fig 2A). Resistance to dicofol is poorly understood. However, target-site mutations that confer resistance to the other acaricides used in this study have been identified (S2 Table). Both MR-VPi and MAR-ABi have the resistance-conferring H110R mutation in METI-Is' target gene *PSST* (*tetur07g05240*) [52,56]. ROS-ITi and MAR-ABi have the two synergistic resistance mutations in the glutamate-gated chloride channel, G314D and G314E in *GluCl1* (*tetur02g04080*) and *GluCl3* (*tetur10g03090*), respectively, and are resistant to the glutamate channel inhibitor abamectin [57]. The L915M mutation in the voltage gated sodium channel (*VGSC*, *tetur34g00970*) is associated with resistance to pyrethroids in *Varroa destructor* [58–60], and was present in all the strains highly resistant to bifenthrin (SOL-BEi, MAR-ABi, and MR-VPi). Further, a recent BSA study that used two sets of fenbutatin oxide susceptible and resistant parental strains, including SOL-BEi and MAR-ABi, identified a candidate target site resistance mutation (V89A) for fenbutatin oxide in the gene encoding subunit c of *mitochondrial ATP-synthase* (*tetur06g03780*) (present in resistant MAR-ABi, absent in susceptible SOL-BEi) [55]. We also found this mutation in ROS-ITi in addition to MAR-ABi; however, as noted by De Beer et al. [55], this gene is duplicated and while the parental MAR-AB outbred strain was reported to have the mutations in both copies, only a single copy bears the mutation after inbreeding in both MAR-ABi and ROS-ITi. This probably explains the absence of resistance in ROS-ITi, and strongly suggests additional metabolic mechanisms of resistance in MAR-ABi. Indeed, and more generally, although the overall presence of target-site resistance mutations is consistent with previous acaricide exposure and can be used as bio-markers [61,62], they nevertheless are not the only determinants of resistance, as multiple studies have now shown [51–53,63]. This is also clear from the data in this study, as resistance ratios varied markedly even among strains with the same target-site resistance mutation (e.g., between MR-VPi and MAR-ABi for two of three METI-I compounds, or between MAR-ABi and ROS-ITi for abamectin), and thus other resistance

mechanisms are almost certainly involved (see Discussion). Finally, although we did not assess resistance of strains to the widely used mite growth inhibitors hexythiazox, clofentezine, and etoxazole, each of the five strains harbored the I1017F target-site variant in *chitin synthase 1* (*CHS-1*, *tetur03g08510*) that confers resistance to these three compounds [64,65].

Fitness on different host plants

We also used fecundity as a proxy for evaluating the performance of four of the inbred strains (ROS-ITi, JP-RRi, MAR-ABi and MR-VPi) on five host plant species known to vary greatly in their production of specialized compounds, as well as other anti-herbivore defenses [39,66]. These included bean (*Phaseolus vulgaris*), a favorable host widely used for the propagation of laboratory strains of *T. urticae*, as well as tomato (*Solanum lycopersicum*), cucumber (*Cucumis sativus*), prune (*Prunus domestica*) and potato (*Solanum tuberosum*). As shown in Fig 2B, the fecundity (as measured by eggs laid) of female mites was higher on bean, cucumber, and prune than on tomato and potato, which belonged to the family Solanaceae. Interestingly, while ROS-ITi performed statistically worse on bean, cucumber, tomato and potato than one or more of the other strains, this pattern was not apparent on prune (ANOVA, with post-hoc t-tests with Tukey corrections for multiple testing, p-values < 0.05). In contrast, JP-RRi showed significantly higher reproductivity on bean, cucumber, prune and potato as compared to one or more other strains. However, it had significantly lower fecundity on tomato than MAR-ABi, with very few eggs laid. MAR-ABi and MR-VPi, which showed the most pronounced multi-acaricide resistance (Fig 2A), did not stand out from other strains on their performance on most hosts. However, on tomato, both MAR-ABi and MR-VPi had higher fecundity than other strains (although only MAR-ABi had significantly higher fecundity than ROS-ITi and JP-RRi). Collectively, our results demonstrate that we captured genetic variation among strains impacting fitness on different host plants.

Variation in gene expression in inbred strains associates imperfectly with genetic distance

Prior transcriptomic studies using strains of *T. urticae* susceptible or resistant to specific acaricides have identified over-expression of detoxification genes in several families as potentially causal for resistance phenotypes [33,53,67–69]. Further, studies using the parental (outbred) populations of MAR-ABi and MR-VPi revealed large differences in the expression of genes in most known detoxification gene families, as well as for many other genes, as compared to the acaricide susceptible London reference strain [33,36]. However, the outbred nature of these strains, coupled with the lack of F1 data, precluded assessment of the regulatory modes for gene expression variation at a genome-wide scale.

To understand the genetic control and inheritance of gene expression variation between *T. urticae* strains that vary in acaricide resistance and host performance, we collected RNA-seq data for each of the five inbred strains included in our study, as well as F1 data from crosses of each of ROS-ITi, JP-RRi, MAR-ABi, and MR-VPi to the SOL-BEi strain. Briefly, SOL-BEi was selected as the common parent in each of the crosses because it was genetically distinct from the other strains (Fig 1A), it was comparatively acaricide sensitive (Fig 2A), and because test crosses revealed that viable F1 progeny were readily produced for each of the four crosses (genetic incompatibilities or partial incompatibilities are not uncommon among *T. urticae* strains [54,70–72]). RNA was collected from 4-7-day old adult females on a single host plant (bean) with either four or five biological replicates for parental strains and resulting F1s (Fig 1B). To assess gene expression variation among genotypes, resulting RNA-seq reads were aligned to the three chromosome London reference genome assembly [51]. To reduce the impact of intra-specific DNA variation among strains on read alignments, genetic variants - identified from genome sequencing and variant prediction (Fig 1A) for each of the five strains - were provided to GSNAP, a variant-aware RNA-seq read mapper (see Methods) [73,74]. A PCA based on transcriptome data revealed that replicates for each strain clustered together and away from other strains, and that F1s had approximately intermediate transcriptomic patterns between each set of parents. While the relative pattern of clustering

of parental expression replicates was broadly similar to that observed for genetic variant data, in the PCA based on expression data JP-RRi clustered more closely to MAR-ABi, and further away from MR-VPi, than observed with genetic data (compare Fig 1B to Fig 1A).

Prominent *trans* control of expression variation among inbred strains

In experimental designs with parental inbred strains (parents) and F1s, the comparison of expression differences between parents, and of ASE in the F1s, allows detection and assignment of regulatory modes underlying gene expression variation. For subsequent analyses, in pairwise comparisons – the crosses of ROS-ITi, JP-RRi, MAR-ABi, and MR-VPi to the SOL-BEi, and the respective F1s – we assessed the genetic mode of control for genes with polymorphisms (SNPs) in exons, a requirement for the measurement of ASE (given variable annotation of untranslated regions in the *T. urticae* genome, we restricted our analyses to coding exons). For consideration for inclusion in a given comparison, we also required genes to be putatively intact in each of the two strains (i.e., no obvious disrupting sequence differences like premature stop codons as assessed from variant predictions), and to have no evidence of copy number variation as assessed with DNA-seq read depth. For these and other criteria used for the selection of the intact+single copy gene sets, including the use of RNA-seq data to further identify potential copy variable genes, see the Methods. Applying these criteria, the resulting numbers of genes for which assignment of genetic mode of control could be attempted ranged from 8,232 in the JP-RRi × SOL-BEi comparison (43.1% of the total 19,087 annotated protein coding genes) to 8,880 in the MR-VPi × SOL-BEi comparison (46.5%); 4,921 genes (25.8%) could be analyzed in all four of the crosses (S1 Fig and S3 Table). The total number of genes, 19,087 in the ~90 Mb *T. urticae* genome, includes many small, hypothetical predictions that were retained in the annotation, a potentially contributing factor to the modest percent of genes included in the analysis (e.g., these sequences may not be functional and therefore conserved). In F1 samples for intact+single copy genes for which SNPs were present, we

recovered and quantified parent-of-origin reads, and quantified reads from the respective parental replicates at the same sites using the same workflow.

By performing differential gene expression analyses between parental strains, and ASE identification in F1s, we determined differential gene expression and assigned modes of control (adjusted $p < 0.1$). After McManus et al. [14], genes showing differential expression either between parents and/or ASE in F1s were sorted into five main categories of control: (1) *cis* effects only (hereafter *cis*), (2) *trans* effects only (hereafter *trans*), (3) synergism, where *cis* and *trans* act in the same direction with respect to one parent (in our case, ROS-ITi, JP-RRi, MAR-ABi, and MR-VPi as compared to SOL-BEi), (4) antagonism, where *cis* and *trans* act in opposite directions with respect to one of the parents, and (5) compensatory, where there is significant ASE but no differential expression between the parents, a result of offsetting *cis* and *trans* effects (a special case of antagonism). Genes showing no differential expression between parental strains and no ASE differences in F1s were classified into a sixth category, conserved. Genes that could not be assigned into any of these six categories because of conflicts between criteria for classification were assigned as ambiguous (further, no attempt at assignment was made for genes with no or low expression, see Methods). In total, genetic mode of control assignments for these categories could be made for between 5,630 (the ROS-ITi × SOL-BEi comparison) to 5,847 (the MAR-ABi × SOL-BEi comparison) of intact+single copy genes (S1 Fig and S4 Table). For 3,381 genes, assignments could be made across all four comparisons.

Across the four comparisons, between 51.9% to 66.4% of genes fell into the *cis*, *trans*, synergism, antagonism, or compensatory categories, with the majority of the remaining genes having conserved expression (S2 Fig and S4 Table). Among all genes for which a genetic mode of control - *cis*, *trans*, synergism, antagonism, or compensatory - could be assigned, those with *trans* control were more numerous by between 2.80- to 3.23-fold as compared to the second most frequent category (*cis*), excepting for the ROS-ITi × SOL-BEi comparison (1.69-fold, Fig 3A-D). However, with increasing magnitude of $\log_2$ fold change ($\log_2$FC) between parents, the relative number of genes classified as regulated in *trans* decreased markedly relative to those regulated in *cis* (e.g., from ~51.4% to ~16.4%

across comparisons in bins of absolute value of parental $\log_2$FC of < 1 to $\log_2$FC >2, while the fraction of genes with *cis* control increased from ~18.1% to ~32.1%, Fig 4A). Further, as fold change increased, genes in the synergism category increased dramatically, an expected finding (i.e., conditioning on increasing fold change enriches for synergism because the two modes of control, *cis* and *trans*, contribute in the same direction to parental fold change). As shown in Fig 3A-D, and in Fig 4A, comparatively little difference in these trends was observed between crosses when considering all genes, apart from the relatively high percent of *cis* control in the lower fold change categories in the ROS-ITi × SOL-BEi comparison (Fig 4A).

Patterns of genetic control for genes associated with xenobiotic resistance and host plant use

Given the phenotypic differences in resistance and host performance among the inbred strains, we examined more specifically differential gene expression and its genetic control for genes in families implicated in the metabolism, transport, and binding of acaricides and plant defense compounds (detoxification gene set, D, 723 genes), host plant use (host genes, H, 318 genes), and a subset of mite peptidases that are potentially involved in overcoming plant produced anti-herbivore protease inhibitors (peptidase gene set, P, 58 genes) (the combined "D+H+P" gene set has 1069 genes; S5 Table). The selection of the detoxification and peptidase gene families and genes are described in the Methods; briefly, the host gene set consists of genes that were previously associated with adaptation of the reference *T. urticae* strain London to tomato as a new host, most of which currently have an unknown function [34] (host genes with overlap to the detoxification gene set were removed from the host genes set, resulting in 288 genes).

Between 32.3-33.2% of genes in the D+H+P gene set were included in our analyses across the four comparisons, a modestly higher fraction than for all genes (29.5-30.6%). As compared to all genes, relatively fewer D+H+P genes fell into the *cis* category in most comparisons, although samples sizes for

D+H+P genes were modest, especially following stratification by the absolute value of $\log_2$FC between parents (compare Fig 4A to Fig 4B). Despite the comparatively small sample sizes, however, a greater relative proportion of D+H+P genes were present in categories of higher absolute value of $\log_2$FC between parents than for the background sets of all genes in each comparison (contrast subpanels between Fig 4A and Fig 4B for each of the four comparisons; all *p*-values for chi-square tests between all genes and D+H+P genes were $< 10^{-5}$). Further, among the D+H+P genes with an absolute value $\log_2$FC > 1 (between parents, or for ASE in F1s), genes in nearly all detoxification gene families, as well as peptidases, were present in each comparison, including a substantial fraction of genes encoding DOGs, lipocalins, and UGTs in one or more comparison (Fig 3E-H). Overall, host genes followed a broadly similar pattern as for the detoxification gene set. Finally, many genes at the extremes of differential expression magnitude belonged to the D+H+P set (compare Fig 3E-H to Fig3A-B), including both up- and downregulated genes (e.g., for genes with an absolute value $\log_2$FC of greater than four, especially for the comparisons of JP-RRi, MAR-ABi, and MR-VPi to SOL-BEi, Fig 3B vs. Fig 3F, Fig 3C vs. Fig 3G, and Fig 3D vs. Fig 3H, respectively).

CYPs linked to metabolism of known acaricides are among the genes with strongest *trans* regulation

The analyses presented in Fig 3 and Fig 4 revealed that many genes in the D+H+P set exhibited strong differential expression, with many controlled in *trans*, some in *cis*, and many with a combination of both. In a complementary approach independent of *a priori* gene family classification, we also investigated the molecular nature of genes with high differential expression as a function of their genetic control mode. To do this, we performed a gene ontology (GO) enrichment analysis (molecular function, MF) on genes showing extremes in fold changes between parents conditioned on contributions of *cis* or *trans* regulation; in this analysis, information across the four comparisons was combined (see Methods), and the top 2.5% of genes in the up- or downregulated tails of the respective distributions were assessed. Strikingly, no GO

terms were significantly enriched when considering *cis* for either up- or downregulation (adjusted $p < 0.05$). However, two GO terms were significantly enriched for *trans*-driven upregulation – GO:0005506 ("iron ion binding") and GO:0016491 ("oxidoreductase activity") (S6 Table). GO:0005506 included eight CYPs and four DOGs. The term GO:0016491 primarily consisted of short-chain dehydrogenases/reductases (SDRs), but also included six other genes, including ones encoding a sorbitol dehydrogenase (*tetur04g07030*) and a phytoene dehydrogenase (*tetur11g04820*). In addition, GO:0005506 ("iron ion binding") was enriched among genes showing high levels of *trans*-driven downregulation, and included six CYPs, three DOGs, and a gene encoding an oxoglutarate/iron-dependent oxygenase (*tetur20g01680*) (S6 Table).

The striking finding of this analysis was that many genes contributing to the *trans* enriched GO terms were already in the detoxification gene set, especially CYPs. Because of this, and because CYPs are important players in detoxification in *T. urticae* and across the animal kingdom [37], we examined more closely the patterns of differential expression and *cis*/*trans* regulation within this gene family (Fig 5). We found that of the eight CYPs enriched for extremes of *trans*-control in the GO analysis (upregulation relative to SOL-BEi, GO:0005506), five belonged to the CYP392 family (Fig 5, CYP2 clan) that has previously been implicated in acaricide detoxification (see Discussion). As compared to SOL-BEi, the most striking upregulation of these genes was observed in the multi-acaricide resistant strains MAR-ABi and MR-VPi, and the least dramatic changes were observed for ROS-ITi, which is comparatively acaricide sensitive. This was particularly striking for the closely related *CYP392A11*, *CYP392A12*, and *CYP392A16* genes, for which strong upregulation was explained predominantly in *trans*. However, *trans*-driven upregulation was not universal for CYP392 family members, as both *CYP392B1* and *CYP392B3* exhibited downregulation resulting from *trans* effects in MAR-ABi and MR-VPi (and for *CYP392B3*, in other strain contrasts as well). For only a few CYP392 genes did *cis* effects contribute substantially to the overall direction of differential expression between parents (i.e., the upregulation of *CYP392E2* in JP-RRi, or the upregulation of *CYP392D1* in all strains). In many cases, synergism or antagonism were

observed, and *cis* effects, albeit relatively minor, contributed synergistically to the very high expression levels of *CYP392A11*, *CYP392A12* and *CYP392A16* in most strains as compared to SOL-BEi.

In addition to CYP392 genes from the CYP2 clan, *trans*-driven upregulation was also observed for *CYP385A1* and *CYP385B1* from the CYP3 clan and *CYP389B1* from the CYP4 clan, and in general a trend for *trans*-driven upregulation of CYP385 genes was observed, particularly in the comparison with ROS-ITi (although $\log_2$FC values were low to modest as compared to some CYP392 genes). Further evidence that related CYP genes can show divergent expression control in *trans* is provided by three CYP389 genes in CYP4 clan with high levels of *trans* control, in which the direction of the *trans* effect was up for *CYP389B1* (for strain MAR-ABi), and down for *CYP389C1* (for strains MAR-ABi and MR-VPi) and *CYP389C11* (for strains JP-RRi, MAR-ABi and MR-VPi).

Although we focused on CYPs, similar patterns were observed for other detoxification genes. For instance, the most striking differences in expression between parents was often explained primarily by *trans* effects, as observed for the DOG genes *DOG1*, *DOG11*, and *DOG13* (in MAR-ABi) (S3 Fig), the lipocalins *tetur24g01030*, *tetur01g05740*, and *tetur01g05730* (in MAR-ABi or MR-VPi) (S4 Fig), or the SDRs *tetur06g05090* (in ROS-ITi and JP-RRi) and *tetur06g04970* (in JP-RRi, MAR-ABi, and MR-VPi) (S5 Fig). Further, as seen for the CYP genes, evidence for regulatory complexity among genotypic comparisons for other detoxification genes was readily apparent (S3-S5 Figs and S4 Table). Examples include the lipocalin gene *tetur01g05740*, which was upregulated in both MAR-ABi and JP-RRi as compared to SOL-BEi, with the former explained predominantly by a *trans* effect, while for the latter only a *cis* effect was predicted. In contrast, for the lipocalin gene *tetur04g06010*, *trans* control contributed to large expression differences in MAR-ABi and MR-VPi as compared to SOL-BEi, but in contrasting directions (up in MAR-ABi, down in MR-VPi), reflecting the opposing parental expression differences.

Dominant inheritance is common for genes controlled in *trans*

In addition to allowing the determination of genetic modes of gene expression control, the parental/F1 experimental design also allowed assignment of expression inheritance, which we assessed by using total read counts for all intact+single copy genes (S3 Table) (note that the determination of inheritance does not rely on ASE). Gene expression inheritance was assigned using the following categories: additive (when F1 expression equals the intermediate expression of parents), dominant (when F1 expression equals the level of either one of the two parents) and transgressive (when F1 expression falls outside the range of both parents) [20]. Using strict criteria, 8.0% to 15.2% of genes were assigned into one of the three inheritance categories across the four comparisons (adjusted $p < 0.01$, and see Methods; S7 Table). Further, when inheritance modes could be assigned, we also recorded when the expression level for parent 1 (P1, here ROS-ITi, JP-RRi, MAR-ABi, or MR-VPi) was greater than for parent 2 (P2, always SOL-BEi), as well as when P1 was less than P2.

For the comparisons with SOL-BEi, where there were significant expression differences between parental strains (either P1>P2 or P1<P2), dominant inheritance was most common, accounting for 68.4%, 72.6% and 76.8% of assignments for MR-VPi, MAR-ABi, and JP-RRi, respectively; in contrast, a lower percentage of genes with dominance was observed in the ROS-ITi comparison (54.6%), for which 43.8% of genes were in the additive category (as compared to 21.6-27.7% for the other three comparisons; Fig 6A and S7 Table). Genes in the transgressive category accounted for a minor proportion of assignments (1.6-3.8%). Additionally, for dominant inheritance there was a bias in the direction of P1 for crosses between JP-RRi, MAR-ABi, and MR-VPi to SOL-BEi (P2) (2.7- to 5.0-fold more than for dominance in the direction of P2), although this trend was not as striking for the ROS-ITi × SOL-BEi comparison (1.3-fold).

We also extended this analysis to specifically examine the D+H+P gene set. Although the resulting number of genes for which inheritance could be assigned was relatively small, broadly similar patterns were observed. Intriguingly, however, a trend for dominance in the direction of the parent with low-

er expression was observed across each of the four comparisons (1.5- to 4.6-fold more genes than for ones with dominance in the direction of the parental line with the higher parental expression level, Fig 6B). We also assessed, for the set of all genes in the additive and dominant categories, the relationship between absolute $\log_2$FC values for genes with significant differences in expression between parents and the relative proportions of additive and dominant inheritance modes. With increasing parental expression differences, the proportion of genes with additive inheritance increased, while the proportion of dominance decreased (Fig 6C). This pattern mirrored what we observed for modes of genetic control, for which *cis* regulation was proportionally greater for genes in bins of increasing magnitude of parental expression differences (categories *cis* and synergism, see Fig 4A). Further, we found that the set of loci featuring *cis* genetic control were enriched for additive inheritance, while *trans* was enriched for dominant inheritance (Fig 6D). A less significant enrichment between *trans* and transgressive inheritance was also observed; in addition, genes controlled by synergism were enriched for additivity, whereas the antagonism category was enriched for transgressive inheritance.

We also examined inheritance patterns for specific detoxification genes with documented or suspected roles in acaricide detoxification in *T. urticae*. Because the inheritance classification we adapted from Bao et al. [20] only assigned complete dominance or additivity, many instances of putative partial dominance were not assigned, including for D+H+P genes. To describe inheritance patterns more fully for these and other genes, we assigned partial dominance and its direction using a post hoc approach (see Methods and S7 Table). Among CYP genes with strong *trans*-driven upregulation in one or more comparisons (Fig 5), *CYP392A11* and *CYP392A12* (synergism with *trans* most prominent in JP-RRi and MR-VPi, and for the latter also in MAR-ABi) showed partial dominance in the direction of the parental line with lower expression (SOL-BEi), as did *CYP392A16* (*trans*) in the comparison involving JP-RRi. In contrast, *CYP392A16* (synergism with a large *trans* effect) in the comparisons with MAR-ABi and MR-VPi exhibited additive inheritance. Incomplete dominance in the direction of P2 (SOL-BEi) was also observed for *CYP385A1* in MAR-ABi (*trans*) and for *CYP385B1* (*trans*) in MR-VPi in comparisons to SOL-BEi. For CYP genes with strong *trans*-driven downregulation in comparison to SOL-BEi, additivity

was observed for *CYP389C11* (synergism) in JP-RRi, and complete dominance in the direction of P1 was observed in MAR-ABi and MR-VPi (*trans*). Many D+H+P genes exhibited complex patterns depending on the comparison. For instance, while *CYP392D3* was expressed more highly in JP-RRi and MR-VPi (P1) as compared to SOL-BEi (P2), additivity was observed for the former, while partial dominance in the direction of P1 was observed for the latter, and for the MAR-ABi to SOL-BEi comparison over-expression relative to both parents was observed in the F1 (a transgressive up classification) (S7 Table).

Discussion

With a parental/F1 experimental design we unraveled the genetic control and inheritance of gene expression variation in strains of *T. urticae*, a highly polyphagous mite pest known for its exceptionally rapid evolution of acaricide resistance [25] and host plant adaptation [30–34]. Key findings were that: (1) *trans*-regulatory variation was common; (2) when expression differences of large magnitude were explained by *trans* effects, detoxification genes, including CYPs and DOGs, were enriched; (3) dominance in gene expression inheritance was often observed, and was enriched among genes with *trans*-regulatory control; and (4) dominance or partial dominance was common for many detoxification genes with dramatic expression differences among strains. It should be noted that while each of the five inbred strains included in our study harbored at least one target-site resistance mutation (a toxicodynamic mechanism), some strains like MAR-ABi had multiple target-site resistance mutations, and also higher resistance to many compounds. However, the relationship between target site mutation numbers and acaricide resistance breadth and levels was incomplete, suggesting an important contribution of acaricide detoxification (metabolic resistance, a toxicokinetic mechanism). Indeed, some target-site mutations in *T. urticae* displayed only moderate phenotypic strength after marker-assisted backcrossing onto a different genetic background [55,63], revealing a requirement for detoxification to attain the very high resistance levels encountered in the field. Further, the strains we used captured variation in host plant use representative of that observed in earlier studies with this species [51,54,75].

Trans regulation is common in *T. urticae*

In all pairwise comparisons, we found that *trans* effects on expression variation were more common than *cis*, even though the latter became more frequent when expression differences between parents were large, as reported in some earlier studies [21,76]. Support for our finding of a higher prevalence of *trans* effects comes as well from inferred modes of inheritance – that is, *trans* effects have been strongly associated with dominance in gene expression ([21], and references therein), and in our study the higher frequency of genes with dominant versus additive inheritance closely mirrored the higher frequency of *trans* versus *cis* effects (this was true among comparisons as well; for instance, for the ROS-ITi x SOL-BEi contrast, relatively fewer *trans* effects and dominance were detected as compared to the other three, compare Fig 3A-D and Fig 4A to Fig 6A). The proportion of *trans* versus *cis* effects has been a focus of many studies, although caution is needed in comparing across studies owing to differences in experimental designs and statistical methods. For instance, our study used more biological replicates than most (five for most comparisons), and more replicates facilitate the detection of differential expression of genes with small fold changes (the group of genes that we found to be most enriched for *trans* effects).

Notwithstanding these considerations, our findings are consistent with previous studies in yeast, plants and animals that have used parental/F1 designs with ASE detection to show that while *cis* effects often predominate in interspecific comparisons, over the shorter evolutionary time frame of intraspecific comparisons, *trans* effects are typically, albeit not always, more common ([22], and references therein). Previously, variation in the performance of *T. urticae* strains on different host plants raised the possibility of host race formation within this species. Despite a striking recent example [54,77], comparatively little evidence exists for wide-spread host-race formation in *T. urticae*, where host plant adaptation can occur in as few as five generations [33], and seasonal or yearly movement of *T. urticae* populations among host plants is likely common. Our finding that *trans* regulation is most often observed in *T. urticae* may reflect this, and contrasts to findings in species like the monkeyflower *Erythranthe guttata* (formerly *Mimulus*

guttatus) and the threespine stickleback fish (*Gasterosteus aculeatus*) for which a predominance of intra-specific *cis* variation has been associated with evolutionary divergent subpopulations (i.e., costal versus inland populations of monkeyflower, and marine versus freshwater populations in the stickleback) [19,78].

Enrichment for *trans* regulation among detoxification genes

Prior studies have revealed large differences in the expression of detoxification genes between *T. urticae* acaricide-susceptible strains and resistant strains that arise rapidly in response to acaricide application [25], or following adaptation to new host plants [30–34]. Selection on *trans*-acting factors is a potential mechanism underlying rapid intraspecific evolution, as many downstream genes impacting the same biological process can be coordinately changed in expression [79,80]. Our study design allowed us to assess this possibility with *T. urticae* strains, including MAR-ABi and MR-VPi, which were derived from outbred progenitor strains that have experienced recent, strong and recurring acaricide selection [33,52,53,69]. Strikingly, while genes in detoxification families were not enriched at the extremes of fold changes driven in *cis*, members of the *CYP* and *DOG* families were enriched at the extremes of *trans*-driven expression variation. At the family level, this was especially striking for *DOG*s, for which the near family-wide upregulation in MAR-ABi compared to SOL-BEi was under *trans*-control with only minor contributions from *cis* effects. Recently, several *T. urticae* DOGs including *DOG11* were shown to have broad substrate specificity against aromatic compounds, including plant-derived specialized compounds with roles in plant defense against herbivores [41]. *DOG11* was the most highly over-expressed *DOG* gene in both MAR-ABi and MR-VPi, both of which had comparatively high performance on tomato. RNAi knockdown of *DOG11* in the MR-VPi progenitor strain was recently found to reduce performance on tomato [41], and whether *trans*-mediated upregulation of *DOG11* contributes to the relatively high fitness of these strains on tomato warrants further study. Apart from CYPs and DOGs, it was also striking that specific genes in most other detoxification families, and some associated with host plant use (host use

or peptidase gene sets), were also among those with the highest differential expression arising from *trans* control. The relatively low frequency of *trans* effects in ROS-ITi, a strain that like SOL-BEi does not exhibit a high level of acaricide resistance, as compared to the other strains – particularly the multi-resistant MR-VPi and MAR-ABi strains – is consistent with selection on *trans*-regulatory variation in response to acaricide selection (although a less clear relationship was observed for JP-RRi). Our findings are therefore generally consistent with a growing number of studies that suggest that variation in *trans* regulation is important for recently evolved pesticide resistance phenotypes, and potentially for host plant adaptation also. For instance, in the mosquito *Aedes aegypti*, three CYPs were overexpressed in the original strain and in a derived introgressed pyrethroid-resistant strain, and were unlinked from the putative, introgressed resistance locus [81].

Notwithstanding findings for the *DOG* family, a robust trend observed across detoxification gene families was that only subsets of genes in families were controlled in *trans*, or that the direction of the expression changes (up or down relative to SOL-BEi) within families or among strains often varied. This finding suggests modular regulation of xenobiotic and host plant associated genes in *T. urticae*. Such modular control might facilitate adaptation to specific host plants and acaricides without wholescale activation of xenobiotic pathways, an outcome that likely entails a fitness cost. In fact, in *T. urticae* evidence for costs to heightened metabolic resistance comes from studies involving *cytochrome P450 reductase* (*CPR*), which encodes the electron donor required for CYP activity. In multiple experimental evolution designs using *T. urticae*, the allele frequencies of *CPR* coding sequence and copy number variants hypothesized to modulate or elevate CPR activity, and that coincided with narrowly-defined intervals for acaricide resistance QTLs, decreased rapidly in the absence of continuous acaricide selection [51–53].

While the parental/F1 experimental approach we used to infer genetic control is a powerful design for detecting *trans*-regulatory variation, it cannot identify the loci that are causal. In fact, both the nature of *trans*-acting factors and the genetic variants that underlie them remain enigmatic. For instance, a recent study of variants affecting *trans* regulation of a reporter gene in yeast, while confirming the role of transcription factors, also noted that 90% were in genes outside the known regulatory network of the tar-

get gene [82]. Given the paucity of information about the regulation of *T. urticae* detoxification pathways, care should be taken in using candidate gene approaches. However, HR96 genes, of which canonical members encode nuclear hormone receptors with ligand-binding domains (LBDs) and DNA-binding domains (DBDs), have been associated with xenobiotic metabolism in insects such as *Drosophila* [83], and are orthologs of the vertebrate xenosensors PXR and CAR [84]. An HR96 transcription factor (*tetur36g00260*) previously associated with the resistant phenotypes of the MAR-ABi and MR-VPi progenitor strains, and to adaptation to tomato for five generations [33], had *trans*-driven upregulation in all four strains in our study (although antagonistic *cis* effects were also observed in MAR-ABi and JP-RRi). Strikingly, while *D. melanogaster* has one HR96 copy, *T. urticae* has eight HR96 homologs (with LBDs and DBDs), and a remarkable 47 genes that have an HR96-like LDB but that lack a DBD [52]. One of the latter (*tetur06g04270*) was at the peak for a resistance QTL for tebufenpyrad resistance in a cross of a strain derived from the MR-VPi progenitor and a susceptible strain [52]. These gene expression and genetic findings, coupled with the exceptional expansion of HR96-like genes in *T. urticae*, raise the possibility that genetic differences at HR96 loci may contribute to the variation in the expression of subsets of detoxification genes observed in our study, although additional work is needed to assess this conjecture.

Multiple modes of genetic control for CYPs known to metabolize acaricides

We observed large expression differences of multiple CYPs known to metabolize acaricides included in our study. This was especially true for the CYP392A clade that includes *CYP392A11*, *CYP392A12*, and *CYP392A16* (CYP2 clan). Especially in multi-acaricide resistant MAR-ABi and MR-VPi, these genes were outliers among all CYPs in their very high expression as compared to SOL-BEi, and were in fact among the most strongly differentially expressed of any genes for which we could assign modes of genetic control. As established in earlier work, CYP392A-clade genes have or are likely to have broad substrate specificity. For example, functional expression studies have shown that CYP392A11 can metabolize fenpyroxymate and cyenopyrafen [85], and CYP392A16 has been documented to metabolize

abamectin [68] as well as pyflubumide [53]. Additionally, RNAi knockdown of CYP392A16 was found to reduce abamectin resistance in the progenitor of the MAR-ABi strain used in our study [69]. For the CYP392A-clade genes in resistant strains as compared to SOL-BEi, upregulation resulted mainly from *trans* effects. Whether the coordinated upregulation of *CYP392A11*, *CYP392A12*, and *CYP392A16* in resistant strains originates from variation in a shared upstream regulator, and whether such a regulator might also control genes in other families, like *DOG* genes for which variation was controlled predominately in *trans* as well, is an outstanding question.

Finally, although *trans* effects were typically largest in magnitude among CYPs, *cis* variation was also observed, including for CYP392A-clade genes in specific pairwise comparisons. In fact, in a prior study promoter variation among two strains, including the MAR-ABi progenitor strain MAR-AB, was demonstrated for *CYP392A16*, with higher expression observed from the MAR-AB promoter than from the promoter of an acaricide-susceptible strain [69]. Adaptive roles for *cis* variants in the control of known detoxification genes in insects have been established or suggested by a number of studies [53,86,87]. For instance, a classic example is the insertion of an Accord transposon into the 5' end of the *Cyp6g1* that confers resistance to DDT in *Drosophila* by elevating *Cyp6g1* expression, and that has spread globally [88]. It was striking that synergism was typically observed where both *trans* and *cis* control of CYP392A-clade genes was inferred. However, a moderately large contribution of a *cis* effect was only seen for *CYP392A11* in strain MR-VPi, where it contributed to the remarkably high expression of *CYP392A11* in this strain as compared to SOL-BEi. The extent to which synergism underpins extremes of expression differences in detoxification gene expression is incompletely understood. Nevertheless, our findings suggest that the evolution of high-level, multi-acaricide resistance phenotypes in response to large and repeated acaricide doses may involve selection on multiple variants acting in a reinforcing manner in both *cis* and *trans*. In particular, synergism leading to exceptionally high expression of detoxification genes such as *CYP392A11* that have broad substrate specificities may have outsized importance. Albeit in a different system, Ji et al. [21] reached a similar conclusion based on their finding that reinforc-

ing *cis* and *trans* effects explained the very high expression of several key and potentially rate-limiting genes in defense pathways in maize lines subjected to intense selection for herbivore resistance.

Dominance or partial dominance is a common inheritance pattern among detoxification genes

The ability of advantageous alleles to respond to selection depends on multiple factors, including the degree of dominance. In diploid individuals, dominant mutations are unmasked to selection in the heterozygous state, and hence selection on dominant variants can occur rapidly even when allele frequencies are low (in populations, low frequency alleles are found predominantly in the heterozygous state). Given the rapid evolution of acaricide resistance in *T. urticae*, an expectation might be that the high expression of genes including those in the CYP392A-clade would result from dominance in the direction of resistant strains like MAR-ABi and MR-VPi. While a trend toward dominance of overall gene expression in the direction of these strains was observed, the direction of dominance for D+H+P genes was instead biased toward the parent with lower expression, even if substantial heterogeneity was observed for specific genes. The significance of this trend, as well as if this could be an effect of the single common parent (SOL-BEi) used in crosses, is currently not clear. More speculatively, it might reflect the action of purifying selection in *T. urticae* populations on variants that in the heterozygous state result in strong upregulation of detoxification and host-associated genes, but that may also have negative pleiotropic effects (i.e., fitness costs of elevated detoxification in the absence of continuous selection by acaricides or challenging host plants). Regardless, for many detoxification and host plant genes, including *CYP392A11*, *CYP392A12*, and *CYP392A16* as well as many *DOG*s that also had very high *trans*-driven expression in the multi-acaricide resistant strains, partial dominance or additivity was observed. The finding of dominance, or the otherwise intermediate expression of many detoxification genes in F1s as compared to parents (i.e., partial dominance or additivity), suggests that genetic variants underlying expression of xenobiotic resistance or host plant use associated genes should be visible to selection in female *T. urticae* mites

(the diploid sex used in our study). This is consistent with the findings of a number of studies that have revealed that phenotypic resistance against pesticides often exhibits incomplete dominance contributed by multiple loci [89–92]. For spider mites, however, an additional factor that likely impacts rapid adaptation is that males are haploid, and in haploid individuals dominance relationships are not relevant. Extrapolation of our findings about gene expression in *T. urticae* females to males will nonetheless require further study.

Conclusion

Among *T. urticae* strains, we found that large changes in the expression of genes known or suspected to be important for detoxification or host plant use were underpinned by variation in *trans*-regulatory environments. This was especially striking for genes encoding detoxification enzymes with known and broad substrate specificities, including for CYPs known to detoxify multiple acaricides, as well as for DOGs that have recently been shown to metabolize a broad range of plant-derived aromatic compounds. Our findings position *T. urticae* as an attractive system for future studies to characterize the molecular basis of the control of gene expression variation in xenobiotic and host plant use pathways.

Methods

Mite strains

Inbred strains used for genetic and gene expression analyses originated from five *T. urticae* strains: ROS-IT, MAR-AB, MR-VP, JP-RR, and SOL-BE. Briefly, SOL-BE was collected on potato in Belgium in 2018, and ROS-IT was collected from roses in Italy in 2017. The MAR-AB strain originated from a heavily sprayed greenhouse near Athens (Greece) in 2009, and is resistant to abamectin, bifenthrin, clofentezine, hexythiazox, fenbutatin oxide and pyridaben [33]. The MR-VP strain originates from a lab stock originally collected in September 2005 from bean plants in a greenhouse in Brussels, Belgium, and was selected for its extraordinary levels of resistance to mitochondrial electron transport inhibitors of

complex I (METI-Is) [48]. The laboratory strain JP-RR was selected for its resistance to electron transport complex II (METI-II) inhibitors (cyanopyrafen, cyflumetofen and pyflubumide) from its parental strain JP-R, as described by Khalighi et al. [49]. To avoid the confounding effect of heterozygosity on downstream studies of the control of gene regulation, the strains were inbred via mother-son mating (performed as described by Bryon et al. [47]) for 7 to 9 generations to produce the respective inbred strains (postfixed with the "i" for inbred designation; ROS-ITi, MAR-ABi, MR-VPi, JP-RRi, and SOL-BEi). The resulting inbred strains were expanded from a single female until the population permitted experimental use. All strains were maintained on potted kidney bean plants (*Phaseolus vulgaris* var. 'Prelude') at 25°C (±0.5°C), 60% relative humidity, and 16:8 h light:dark photoperiod.

Bioassays

Acaricide resistance levels of all inbred strains were determined in bioassays as previously described by Van Pottelberge et al. [48]. Serial dilutions of acaricides and water (negative control) were tested in four replicates of 20-30 adult female mites on 9 cm^2 square leaf disks on wet cotton. Using a Cornelis spray tower, 800 µl of fluid was sprayed on each leaf-disk at a pressure of 1 bar (1.5 mg fluid deposition per cm^2). The acaricides used and the corresponding IRAC Mode of Action (MoA) and time of mortality assessment are listed in S1 Table. Dose-response curves, LC_{50}-values, and their 95% confidence intervals were calculated using POLO-plus software (version 2.0) [93]; after earlier work, significant differences were assessed when 95% confidence intervals did not overlap [94].

Host performance assays

The fitness of inbred strains ROS-ITi, JP-RRi, MAR-ABi, and MR-VPi was assessed on five different hosts: bean (*Phaseolus vulgaris* var. 'Prelude'), tomato (*Solanum lycopersum* var. 'Moneymaker'), potato (*Solanum tuberosum* var. 'Bintje'), cucumber (*Cucumis sativus* var. 'Ventura') and prune (*Prunus domestica* var. 'Brompton'). This was evaluated by placing leaf discs of the respective host plant (4 cm^2)

on wet cotton wool with the adaxial side up. Discs were infested with ten female adults (4-5 days old) per disc with at least four replicates per strain and per host. After three days, the total number of eggs on each leaf disc was recorded. Statistical differences among strains by host plant were assessed by analysis of variance (ANOVA); where significant differences were observed, post-hoc Tukey tests were applied. Analyses were performed in R (version 4.0.4) [95].

DNA extractions and sequencing

For each of the parental inbred strains, DNA was isolated from a pool of 400 adult female mites by phenol-chloroform extraction as described in Villacis-Perez et al. [54]. DNA concentration and purity were measured using a spectrophotometer (DeNovix, USA). Integrity of the DNA samples was checked via gel electrophoresis (2% agarose gel; 30 min; 100V). From the resulting DNA samples of the MAR-ABi, MR-VPi, JP-RRi, and SOL-BEi strains, Illumina DNA-seq libraries were constructed using the Illumina TruSeq Nano DNA library prep kit, and subsequently sequenced with the Illumina HiSeq2500 technology (paired-end reads of 125 bp) at the Huntsman Cancer Institute of the University of Utah (Salt Lake City, Utah, USA). From the DNA extracted for ROS-ITi, an Illumina library was constructed using the Illumina NEBNext Ultra II DNA library prep kit and sequenced with the Illumina HiSeq3000 technology (paired-end reads of 150 bp) at NXTGNT (Ghent, Belgium).

DNA read alignments and PCA

Genomic reads from each of the five inbred strains were aligned to the *T. urticae* London reference genome [51] using BWA (version 0.7.17-r1188) [96]. Sorting and indexing of the resulting BAM files were performed with SAMtools (version 1.9) [97], and duplicate reads were marked with Picard (version 2.18.11-SNAPSHOT) (http://broadinstitute.github.io/picard/). Genetic variants across all strains were predicted with the Genomic Analysis Toolkit (GATK) following the best practices workflow (version 4.0.7.0) [98,99]. To assess potential acaricide resistance target-site changes, potential effects of predicted

variants on amino acid sequences were assessed using a custom-built database with SNPeff (version 4.3L) [100]. The database was constructed based on the coding exons (CDS) from the June 2016 ORCAE GFF annotation [101]. To describe relatedness among strains, a principal component analysis (PCA) was performed after filtering the resultant variant call format (VCF) file with additional criteria to further eliminate aberrant predictions. The filtering criteria were adapted from previous studies [51]. Briefly, each SNP included in the PCA had to have the quality score normalized by allele depth (QD in the VCF file) of at least 2, mean mapping quality score (MQ) of at least 50, and strand odds ratio (SOR) below 3. Mapping quality rank sum (MQRankSum) and rank sum for relative positioning of alleles in reads (ReadPosRankSumTest) both had to be at least −8 when present. Finally, coverage (as indicated by the total depth per allele per sample [AD] field in the VCF file) had to be within 25% and 150% of the sample's genome-wide mean SNP read coverage (to restrict the analysis to variants in putative single copy regions). The PCA was carried out in R (version 3.5.2) [95] using SNPRelate (version 0.16.1) [102] with the "autosome.only=FALSE" option. The PCA was performed on the GDS file converted from the VCF using gdsfmt (version 0.18.1) [102,103] with the option "method=copy.num.of.ref" (which includes all SNPs). Further, we confirmed that the five strains were in fact isogenic, as expected for inbred strains, by determining the fraction of variable SNPs in each parental strain that were fixed (i.e., homozygous; see Results); to do this, we used the approach of Villacis-Perez et al. [54].

Genetic crosses for genetic mode of expression control determination

To investigate expression variation between strains and its genetic mode of control and inheritance, we adapted a parental/F1 design that has been used previously [9,14,15,20,104]. To do this, we crossed 40 virgin females of strains ROS-ITi, JP-RRi, MAR-ABi, and MR-VPi to 40 males of strain SOL-BEi to generate four sets of F1s; in parallel, we produced matching progeny for each parental strain. The crosses to generate F1 and parental strain females were performed with 5-fold biological replication, except for MAR-ABi, JP-RRi, and the JP-RRi × SOL-BEi F1s, for which four biological replicates were obtained.

For each replicate (F1 or parental), 100-120 adult female mites at four days of age were collected, except for the JP-RRi × SOL-BEi F1s for which females of 4-7 days old were collected. Immediately after mite collection, samples were frozen in dry ice until storage at -80°C.

RNA extractions and sequencing

For each sample of 100-120 adult female mites (a parental or F1 replicate), total RNA was extracted from the frozen mite samples (-80°C) using the RNeasy Plus Mini Kit (Qiagen, Belgium) according to the manufacturer's Quick-Start Protocol. Samples were kept on dry ice until adding the first buffer of the RNeasy Plus Mini Kit. RNA concentration and purity were measured using a DeNovix spectrophotometer (DeNovix, USA). The integrity of the RNA-sample was checked via gel electrophoresis (1% agarose gel; 30 min; 100 V). Illumina libraries were constructed using the Illumina TruSeq stranded mRNA library prep kit and sequenced on the Illumina HiSeq3000 platform (paired-end reads of 150 bp). Library construction and sequencing was conducted at NXTGNT (Ghent, Belgium).

RNA-seq read alignments and PCA

RNA-seq reads from parents and F1 samples were aligned to the London reference genome using GSNAP (version 2018-07-04) [73,74], an aligner that incorporates SNP information to reduce reference bias in the alignment of reads derived from non-reference samples. The alignments were performed with known splice sites supplied to GSNAP using the GFF annotation of Wybouw et al. [51] with novel splice discovery enabled and with the options "--localsplicedist=50000 --novelend-splicedist=50000 --pairmax-rna=50000 --gmap-mode=all". Variant information was supplied to GSNAP in a sample-dependent manner. Briefly, pairwise VCF files were generated and supplied to GSNAP for the respective alignments. For instance, for alignments of reads from F1s from the ROS-ITi × SOL-BEi cross, a VCF file was supplied to GSNAP that contained only the SNP variants identified for the ROS-ITi and SOL-BEi strains as compared to the London reference genome (see section "DNA read alignments and PCA"). Read counts

per gene per genotype and biological replicate were then assessed with HTSeq version 0.11.2 [105]. To assess global patterns of gene expression, as well as variation among replicates for all parental strain and F1 samples, a PCA was performed on the resulting expression data using DESeq2 (version 1.22.2) [106].

Selection of putatively intact and single copy gene sets

For pairwise strain expression comparisons to detect ASE (i.e., for each set of two parent strains and the respective F1), we first identified genes that were putatively intact and single copy in each pair of parents. To select the "intact+single copy" gene sets, we used both genomic read and RNA-seq read data. With aligned genomic reads on a per strain basis, we detected putative copy number variation using read coverage. To do this, median per-base DNA read coverage was determined across the three *T. urticae* chromosomes. Briefly, pysam's "count coverage" function was used with default settings (pysam version 0.15; https://pysam.readthedocs.io/en/latest/api.html), and the median of all non-zero values was recovered. In addition, the coverage for each CDS exon of a gene was assessed using the same approach with reads that mapped in a proper pair and with a MQ score of 40 (the maximum quality score). Genes that had a coverage below 5% or above 150% of the genome-wide median within any CDS exon were considered to be copy variable in the respective strain. In addition, GATK variant predictions (single nucleotide and small indel differences) were assessed for impacts on genes by substituting them into the London reference sequence on a per strain basis. The respective CDS sequences for each gene were then extracted and joined to determine the largest open reading frame (ORF) accounting for the strand on which the gene was annotated ("+" or "-" in the GFF annotation file). Biopython version 1.74 was then used for translations [107]. To account for variation at gene ends that may be neutral, such as nearby alternative start or stop codons, sequences 12 bp upstream of the first CDS exon and 12 bp downstream of the last CDS exon were included in the ORF prediction.

In order for a gene to be considered for inclusion in downstream ASE analyses involving a pair of strains, based on genomic read data the following criteria had to be met: (1) the gene had to be putatively

single copy in both strains as assessed from DNA read coverage, (2) the longest ORF of the gene had to contain a putative start and stop codon in both strains, and (3) there had to be an amino acid sequence match of at least 90% for the predicted protein sequences between the two strains. For assessing the latter, global alignments were performed using the pairwise2 module in Biopython; identical amino acids were given an assigned a score of 1, and the sum of scores was divided by the alignment length (the longer of the two sequences used in each pairwise comparison) to determine percent identity.

Additionally, to be considered in an ASE analysis between a pair of parents, a gene had to pass a final filter based on RNA-seq read alignment data. Briefly, apparent heterozygosity in aligned RNA-seq reads, if observed in inbred strains, may result from cross-mapping of similar, duplicate sequences to a single sequence in the reference genome used for alignment. To detect this, for each inbred strain we combined RNA-seq reads from all replicates (i.e., from all BAM files for a given parental genotype) and selected reads in proper pairs with MQ values of 40. For each position within a CDS, we required that the minor allele frequency could be at most 0.1 (in assessing base calls in reads by position, a minimum base quality score threshold of 30 was used). If a gene failed this criterium in either of the two parents, it was not included in the respective intact+single copy gene set.

Assignment of parent of origin to aligned RNA-seq reads

ASE analyses based on RNA-seq alignments require that individual reads are assigned to a parent of origin based on variants in transcribed regions (reads that align to identical sequences between the two parent strains are uninformative). For F1 samples, we developed an in-house Python program (available at https://github.com/akurlovs/ase) that employed pysam (version 0.15) to distinguish reads derived from one or the other parent for a respective F1 sample. Briefly, SNPs that distinguished the two parents of a given F1 were first recovered from the respective pairwise VCF files (see section “RNA-seq read alignments and PCA”; an MQ value of 40 or more was applied for SNP selection). F1 RNA-seq reads that were mapped in a proper pair with read-level MQ values of 40 (the maximum values assigned with

GSNAP) were then recovered and written to parent of origin BAM files if reads contained bases at SNP sites specific to one of the two parents of the F1 (reads for which the parent of origin could not be assigned were discarded). The same workflow was also applied to RNA-seq samples of the respective parents to generate matching read alignment data (that is, only parental reads that overlapped the same informative SNP sites used in the F1 analyses were included). Following parent of origin read assignment, reads within CDS exons were counted from the resulting BAM files with HTSeq.

Assignment of mode of genetic control for gene expression variation

By assessing differential expression between parent of origin expression in F1s (ASE, *cis* effects only) and differential expression between the parents (which can arise from both *cis* and *trans* effects), we classified genes into seven control modes adapted from McManus et al. [14]: "*cis*", "*trans*", "synergism" (*cis* and *trans* effects acting in the same direction), "antagonism" (*cis* and *trans* effects acting in opposite directions), "compensatory" (a special case of antagonism for which significant *cis* and *trans* effects offset such that there is no differential expression between parents), "conserved" (no parental expression differences and no significant *cis* or *trans* effects), and "ambiguous" (genes that do not fall in any of the above categories). Differential expression analyses between parents of an F1 and between parent-of-origin reads for the respective F1s were both performed with DESeq2 with the alpha value set to the default of 0.1 (i.e., an adjusted *p*-value threshold of 0.1 was used). Genes with average base mean expression values of below 20, as determined in the respective DESeq2 analyses, were excluded from subsequent analyses for assignment of genetic mode of control. We then assigned significant *trans* effects by applying Fisher's exact test to DESeq2 normalized allele-specific count ratios of F1 versus normalized count ratios of respective parental strains for genes in "intact+single copy" gene sets by pairwise comparison; for a given cross, *p*-values for *trans* assignment were then Benjamini-Hochberg adjusted for multiple comparisons and those for which the adjusted-*p* was less than 0.1 were considered significant. For performing tests for significant *trans* effects, the DESeq2 normalized expression data across replicates was summed, and re-

sulting values were weighted by the number of replicates if different numbers of replicates were available for given pairwise comparisons. Fold changes between parents, as well as for F1s (*cis* effects), were from estimates from the DESeq2 analyses ($\log_2$FC values); the contribution of *trans* effects to fold changes was estimated by subtracting the respective $\log_2$FC values for the F1 parent-specific allele counts (ASE analyses) from the $\log_2$FC values for the two parental strains. Further, prior to calculating $\log_2$FC values associated with *trans* regulation, those for the parental and F1s analyses were shrunk using ashr (version 2.2-32) [108] to prevent lowly expressed genes from confounding the results.

Inheritance of gene expression variation

By comparing expression of genes between the parental strains and that of respective hybrid F1 offspring, the inheritance mode for expression of individual genes was inferred. For this purpose, we adapted the methodology from Bao et al. [20]. The inheritance modes were classified into the following categories: "additive" (expression in F1s is equally between that of the two parental strains), "dominant" (F1 expression is equal to that of either the maternal or paternal strain), and "transgressive" (F1 expression falling outside the parental range). Note that while additivity and dominance are observed when there are significant expression differences between parental strains, transgressive inheritance can be observed even if there is no expression divergence between parents. For this inheritance mode classification, we performed four rounds of differential gene expression analyses using DESeq2 (version 1.28). Input for the analyses was total gene read counts for all genes (all replicates were used for every genotype, parental or F1); for the following, P1 represents the varying parents (ROS-ITi, JP-RRi, MAR-ABi, and MR-VPi) and P2 represents the SOL-BEi strain. The first differential expression analysis was performed using the parents, and significantly differentially expressed genes between parents (P_1 $\overline{P}_2$) were identified with the following criteria: (1) *p*-adjusted < 0.01, (2) absolute $\log_2$FC > 0.5, and (3) baseMean > 15 (to remove lowly expressed genes from the analysis). The second and third rounds of analyses were performed by comparing F1 expression to the respective two parental strains using a significance cutoff of *p*-adjusted < 0.01. The

fourth analysis was performed by comparing F1 expression to the midpoint of parental strains (*p*-adjusted < 0.01; see Bao et al. [20] for methods). In the case of P1 $\overline{P}$2, if F1=P1 or F1=P2 only, genes were classified into the dominant category in the direction of one or the other parent, and for F1=(P1+P2)/2 only, genes were classified into the additive inheritance category. The criteria for transgressive inheritance was (1) F1>P1 and F1>P2, or (2) F1<P1 and F1<P2.

The strict criteria we used to assign genes into pure dominant, additive, and transgressive categories overlooked a large number of genes showing partial dominance when P_1 $\overline{P}_2$. To more fully describe inheritance, we used the formula of Stone [109] to assign partial dominance and its direction:

$$\mathrm{H} = \frac{2F1 - P1 - P2}{P1 - P2}$$

where F1, P1 and P2 represent respective mean expression values of each genotype (after normalization by DESeq2). When 0 < H < 1, incomplete dominant in the direction of P1 was assigned; when -1 < H < 0, incomplete dominant in the direction of P2 was assigned.

Association analysis between inheritance mode (excluding partial dominance categories) and mode of gene regulatory control was based on Pearson's Chi-squared test using the chisq.test function in R v4.0 [95] with data across all four comparisons for the subset of genes where both inheritance mode and gene regulatory modes were assigned.

Detoxification and host plant associated genes

"Detoxification genes" were defined as genes belonging to gene families involved in detoxification and/or host plant adaptation, although it should be noted that not every gene in these families plays a role in the detoxification process. For ATP-binding cassette transporters (ABCs), carboxyl/cholinesterases (CCEs), glutathione S-transferases (GSTs), intradiol ring-cleavage dioxygenases (DOGs), and UDP-glycosyltransferases (UGTs) gene family information was obtained directly from the *T. urticae* three-chromosome gene information file [51]. Since the release of the 2016 annotation, Snoeck et al. [46] reannotated short-chain dehydrogenases (SDRs) and PLAT-domain proteins in *T. urtica*e and gene identi-

fiers for those families were obtained from "Tables *S7* and *S8*" in their manuscript (only complete sequences were used). In addition, genes belonging to the cytochrome P450 (CYP) family and the major facilitator superfamily (MFS) were identified using InterProScan (version 5.32-71.0) [110], which was run on the peptides specified in the June 2016 ORCAE annotation [101], by extracting genes annotated as "IPR036396" ("Cytochrome P450 superfamily") and "IPR011701" ("Major facilitator superfamily", Pfam category identified as upregulated upon host acclimation in Snoeck et al. [46]), respectively. The lipocalin gene set came from "Table S2" in Dermauw et al. [33].

The 318 "core" genes that were highly and consistently differentially expressed upon host transfer in Wybouw et al. [34] were also included and referred to as "host" genes. The list of peptidases was obtained from Grbić' et al.'s [42] supplementary list "S6.1.9. Cysteine peptidase genes" and "S6.1.10. Aspartic peptidase genes", except for *tetur13g02490* which was replaced by *tetur13g02500* (a name change in later annotations). See S5 Table for resulting family and host classifications for genes.

Gene ontology (GO) enrichment analysis

Gene ontology enrichment analyses (molecular function, MF) for genes showing extreme upregulating or downregulating *cis* and *trans* effects – defined as those in the top 2.5% and bottom 2.5% for *cis* and *trans* evaluated independently – were performed on the GO categories annotated on the June 2016 *T. urticae* ORCAE annotation [101] by InterProScan (version 5.32-71.0) [110]. For the pairwise comparisons to SOL-BEi, top indicates upregulation relative to SOL-BEi, and bottom indicates downregulation relative to SOL-BEi. The distributions for *cis* and *trans* were established based on the maximum contributions of *cis* or *trans* for given genes assessed across all four comparisons; where a gene was controlled by combinations of *cis* and *trans* (e.g., control by synergism), the contributions of each of *cis* and *trans* were considered independently in constructing the respective distributions. For GO categories that included ten or more genes in the sample being tested, we determined whether the tails of the distributions contained a disproportionately high number of genes belonging to the GO category by performing Fisher's exact tests

(the comparison was to all genes in the respective distributions). The resulting *p*-values were Benjamini-Hochberg-corrected to adjust for multiple comparisons and enriched GO terms were identified where the adjusted *p*-values were < 0.05.

Phylogenetic constructions of detoxification gene families

Neighbor-joining trees were constructed for the CYP, DOG, lipocalin, and SDR detoxification gene families. Amino acid sequences for the respective families were extracted from the June 2016 annotation [101] and aligned using MUSCLE (version 3.8.1551) [111]. The scikit-bio Python package (version 0.5.4) (https://github.com/biocore/scikit-bio) was used to construct neighbor-joining trees based on a distance matrix between the amino acid sequences. The trees were visualized alongside heatmaps displaying differential expression magnitude ($\log_2$FC) between parental strains, as well as expression explained by *cis* and *trans* effects, using the Python package ete3 (version 3.1.1) [112] and matplotlib [113]. For display, a gene was included in a given pairwise comparison if a genetic mode of control could be assessed.

Acknowledgements

We thank Ibrahim Ismaeil for help with inbreeding and bioassays, and Robert Greenhalgh for assistance with managing read data. This work was supported by grant 772026-POLYADAPT, 773902–SuperPests (Research Council (ERC) under the European Union's Horizon 2020 research and innovation program) and BOFSTA2017003701 (Bijzonder Onderzoeksfonds UGent) to TVL, and by incentive funds from the University of Utah to RMC.

Data availability statement

DNA and RNA-seq read data have been deposited to NCBI under identifiers PRJNA799176 and PRJNA801103, respectively. Pairwise strain variant call data, and read count data, have been deposited to figshare (https://doi.org/10.6084/m9.figshare.20291175.v1).

Competing interests

The authors declare no competing interests.

References

1. Slatkin M. Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. Nat Rev Genet. 2008;9: 477–485. doi:10.1038/nrg2361

2. Charlesworth B. Effective population size and patterns of molecular evolution and variation. Nat Rev Genet. 2009;10: 195–205. doi:10.1038/nrg2526

3. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. Nat Rev Genet. 2006;7: 85–97. doi:10.1038/nrg1767

4. Alkan C, Coe BP, Eichler EE. Genome structural variation discovery and genotyping. Nat Rev Genet. 2011;12: 363–376. doi:10.1038/nrg2958

5. Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, et al. *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. Science. 2000;289: 85–88. doi:10.1126/science.289.5476.85

6. Shapiro MD, Marks ME, Peichel CL, Blackman BK, Nereng KS, Jónsson B, et al. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature. 2004;428: 717–723. doi:10.1038/nature02415

7. Gompel N, Prud'homme B, Wittkopp PJ, Kassner VA, Carroll SB. Chance caught on the wing: *cis*-regulatory evolution and the origin of pigment patterns in *Drosophila*. Nature. 2005;433: 481–487. doi:10.1038/nature03235

8. Clark RM, Wagler TN, Quijada P, Doebley J. A distant upstream enhancer at the maize domestication gene *tb1* has pleiotropic effects on plant and inflorescent architecture. Nat Genet. 2006;38: 594–597. doi:10.1038/ng1784

9. Wittkopp PJ, Haerum BK, Clark AG. Evolutionary changes in *cis* and *trans* gene regulation. Nature. 2004;430: 85–88. doi:10.1038/nature02698

10. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009;10: 57–63. doi:10.1038/nrg2484

11. Albert FW, Kruglyak L. The role of regulatory variation in complex traits and disease. Nat Rev Genet. 2015;16: 197–212. doi:10.1038/nrg3891

12. Landry CR, Wittkopp PJ, Taubes CH, Ranz JM, Clark AG, Hartl DL. Compensatory *cis-trans* evolution and the dysregulation of gene expression in interspecific hybrids of *Drosophila*. Genetics. 2005;171: 1813–1822. doi:10.1534/genetics.105.047449

13. Tirosh I, Reikhav S, Levy AA, Barkai N. A yeast hybrid provides insight into the evolution of gene expression regulation. Science. 2009;324: 659–662. doi:10.1126/science.1169766

14. McManus CJ, Coolon JD, Duff MO, Eipper-Mains J, Graveley BR, Wittkopp PJ. Regulatory divergence in *Drosophila* revealed by mRNA-seq. Genome Res. 2010;20: 816–825. doi:10.1101/gr.102491.109

15. Shi X, Ng DW-K, Zhang C, Comai L, Ye W, Jeffrey Chen Z. *Cis*- and *trans*-regulatory divergence between progenitor species determines gene-expression novelty in *Arabidopsis allopolyploids*. Nat Commun. 2012;3: 950. doi:10.1038/ncomms1954

16. Schaefke B, Emerson JJ, Wang T-Y, Lu M-YJ, Hsieh L-C, Li W-H. Inheritance of gene expression level and selective constraints on *trans*- and *cis*-regulatory changes in yeast. Mol Biol Evol. 2013;30: 2121–2133. doi:10.1093/molbev/mst114

17. Waters AJ, Makarevitch I, Noshay J, Burghardt LT, Hirsch CN, Hirsch CD, et al. Natural variation for gene expression responses to abiotic stress in maize. Plant J. 2017;89: 706–717. doi:10.1111/tpj.13414

18. Glaser-Schmitt A, Zečić A, Parsch J. Gene regulatory variation in *Drosophila melanogaster* renal tissue. Genetics. 2018;210: 287–301. doi:10.1534/genetics.118.301073

19. Verta J-P, Jones FC. Predominance of *cis*-regulatory changes in parallel expression divergence of sticklebacks. Elife. 2019;8: e43785. doi:10.7554/eLife.43785

20. Bao Y, Hu G, Grover CE, Conover J, Yuan D, Wendel JF. Unraveling *cis* and *trans* regulatory evolution during cotton domestication. Nat Commun. 2019;10: 5399. doi:10.1038/s41467-019-13386-w

21. Ji M, Bui H, Ramirez RA, Clark RM. Concerted *cis* and *trans* effects underpin heightened defense gene expression in multi-herbivore-resistant maize lines. Plant J. 2022. doi:10.1111/tpj.15812

22. Hill MS, Vande Zande P, Wittkopp PJ. Molecular and evolutionary processes generating variation in gene expression. Nat Rev Genet. 2021;22: 203–215. doi:10.1038/s41576-020-00304-w

23. Van Leeuwen T, Vontas J, Tsagkarakou A, Dermauw W, Tirry L. Acaricide resistance mechanisms in the two-spotted spider mite *Tetranychus urticae* and other important Acari: A review. Insect Biochem Mol Biol. 2010;40: 563–572. doi:10.1016/j.ibmb.2010.05.008

24. Feyereisen R, Dermauw W, Van Leeuwen T. Genotype to phenotype, the molecular and physiological dimensions of resistance in arthropods. Pestic Biochem Physiol. 2015;121: 61–77. doi:10.1016/j.pestbp.2015.01.004

25. Van Leeuwen T, Dermauw W. The molecular evolution of xenobiotic metabolism and resistance in chelicerate mites. Annu Rev Entomol. 2016;61: 475–498. doi:10.1146/annurev-ento-010715-023907

26. Hawkins NJ, Bass C, Dixon A, Neve P. The evolutionary origins of pesticide resistance. Biol Rev. 2019;94: 135–155. doi:10.1111/brv.12440

27. Roush RT, McKenzie JA. Ecological genetics of insecticide and acaricide resistance. Annu Rev Entomol. 1987;32: 361–380. doi:10.1146/annurev.en.32.010187.002045

28. Nkongolo KK, Quick JS, Limin AE, Fowler DB. Sources and inheritance of resistance to Russian wheat aphid in *Triticum* species amphiploids and *Triticum tauschii*. Can J Plant Sci. 1991;71: 703–708. doi:10.4141/cjps91-103

29. Paolino A, Gassmann A. Assessment of inheritance and fitness costs associated with field-evolved resistance to Cry3Bb1 maize by Western corn rootworm. Toxins. 2017;9: 159. doi:10.3390/toxins9050159

30. Gould F. Rapid host range evolution in a population of the phytophagous mite *Tetranychus urticae* Koch. Evolution. 1979;33: 791. doi:10.2307/2407646

31. Fry JD. Evolutionary adaptation to host plants in a laboratory population of the phytophagous mite *Tetranychus urticae* Koch. Oecologia. 1989;81: 559–565.

32. Magalhães S, Fayard J, Janssen A, Carbonell D, Olivieri I. Adaptation in a spider mite population after long-term evolution on a single host plant: Host plant adaptation in spider mites. J Evol Biol. 2007;20: 2016–2027. doi:10.1111/j.1420-9101.2007.01365.x

33. Dermauw W, Wybouw N, Rombauts S, Menten B, Vontas J, Grbić M, et al. A link between host plant adaptation and pesticide resistance in the polyphagous spider mite *Tetranychus urticae*. Proc Natl Acad Sci USA. 2013;110: E113–E122. doi:10.1073/pnas.1213214110

34. Wybouw N, Zhurov V, Martel C, Bruinsma KA, Hendrickx F, Grbić V, et al. Adaptation of a polyphagous herbivore to a novel host plant extensively shapes the transcriptome of herbivore and host. Mol Ecol. 2015;24: 4647–4663. doi:10.1111/mec.13330

35. Krieger RI, Feeny PP, Wilkinson CF. Detoxication enzymes in the guts of caterpillars: an evolutionary answer to plant defenses? Science. 1971;172: 579–581. doi:10.1126/science.172.3983.579

36. Dermauw W, Osborne EJ, Clark RM, Grbić M, Tirry L, Van Leeuwen T. A burst of ABC genes in the genome of the polyphagous spider mite *Tetranychus urticae*. BMC Genomics. 2013;14: 317. doi:10.1186/1471-2164-14-317

37. Dermauw W, Van Leeuwen T, Feyereisen R. Diversity and evolution of the P450 family in arthropods. Insect Biochem Mol Biol. 2020;127: 103490. doi:10.1016/j.ibmb.2020.103490

38. Gloss AD, Vassão DG, Hailey AL, Nelson Dittrich AC, Schramm K, Reichelt M, et al. Evolution in an ancient detoxification pathway is coupled with a transition to herbivory in the Drosophilidae. Mol Biol Evol. 2014;31: 2441–2456. doi:10.1093/molbev/msu201

39. Heidel-Fischer HM, Vogel H. Molecular mechanisms of insect adaptation to plant secondary compounds. Curr Opin Insect Sci. 2015;8: 8–14. doi:10.1016/j.cois.2015.02.004

40. Pavlidi N, Vontas J, Van Leeuwen T. The role of glutathione S-transferases (GSTs) in insecticide resistance in crop pests and disease vectors. Curr Opin Insect Sci. 2018;27: 97–102. doi:10.1016/j.cois.2018.04.007

41. Njiru C, Xue W, De Rouck S, Alba JM, Kant MR, Chruszcz M, et al. Intradiol ring cleavage dioxygenases from herbivorous spider mites as a new detoxification enzyme family in animals. BMC Biol. 2022;20: 131. doi:10.1186/s12915-022-01323-1

42. Grbić M, Van Leeuwen T, Clark RM, Rombauts S, Rouzé P, Grbić V, et al. The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. Nature. 2011;479: 487–492. doi:10.1038/nature10640

43. Wybouw N, Van Leeuwen T, Dermauw W. A massive incorporation of microbial genes into the genome of *Tetranychus urticae*, a polyphagous arthropod herbivore. Insect Mol Biol. 2018;27: 333–351. doi:10.1111/imb.12374

44. Ahn S-J, Dermauw W, Wybouw N, Heckel DG, Van Leeuwen T. Bacterial origin of a diverse family of UDP-glycosyltransferase genes in the *Tetranychus urticae* genome. Insect Biochem Mol Biol. 2014;50: 43–57. doi:10.1016/j.ibmb.2014.04.003

45. Wybouw N, Pauchet Y, Heckel DG, Van Leeuwen T. Horizontal gene transfer contributes to the evolution of arthropod herbivory. Genome Biol Evol. 2016;8: 1785–1801. doi:10.1093/gbe/evw119

46. Snoeck S, Wybouw N, Van Leeuwen T, Dermauw W. Transcriptomic plasticity in the arthropod generalist *Tetranychus urticae* upon long-term acclimation to different host plants. G3. 2018; g3.200585.2018. doi:10.1534/g3.118.200585

47. Bryon A, Kurlovs AH, Dermauw W, Greenhalgh R, Riga M, Grbić M, et al. Disruption of a horizontally transferred phytoene desaturase abolishes carotenoid accumulation and diapause in *Tetranychus urticae*. Proc Natl Acad Sci USA. 2017;114: E5871–E5880. doi:10.1073/pnas.1706865114

48. Van Pottelberge S, Van Leeuwen T, Nauen R, Tirry L. Resistance mechanisms to mitochondrial electron transport inhibitors in a field-collected strain of *Tetranychus urticae* Koch (Acari: Tetranychidae). Bull Entomol Res. 2009;99: 23. doi:10.1017/S0007485308006081

49. Khalighi M, Dermauw W, Wybouw N, Bajda S, Osakabe M, Tirry L, et al. Molecular analysis of cyenopyrafen resistance in the two-spotted spider mite *Tetranychus urticae*. Pest Manag Sci. 2016;72: 103–112. doi:10.1002/ps.4071

50. Xue W, Snoeck S, Njiru C, Inak E, Dermauw W, Van Leeuwen T. Geographical distribution and molecular insights into abamectin and milbemectin cross resistance in European field populations of *Tetranychus urticae*. Pest Manag Sci. 2020;76: 2569–2581. doi:10.1002/ps.5831

51. Wybouw N, Kosterlitz O, Kurlovs AH, Bajda S, Greenhalgh R, Snoeck S, et al. Long-term population studies uncover the genome structure and genetic basis of xenobiotic and host plant adaptation in the herbivore *Tetranychus urticae*. Genetics. 2019;211: 1409–1427. doi:10.1534/genetics.118.301803

52. Snoeck S, Kurlovs AH, Bajda S, Feyereisen R, Greenhalgh R, Villacis-Perez E, et al. High-resolution QTL mapping in *Tetranychus urticae* reveals acaricide-specific responses and common target-site resistance after selection by different METI-I acaricides. Insect Biochem Mol Biol. 2019;110: 19–33. doi:10.1016/j.ibmb.2019.04.011

53. Fotoukkiaii SM, Wybouw N, Kurlovs AH, Tsakireli D, Pergantis SA, Clark RM, et al. High-resolution genetic mapping reveals *cis*-regulatory and copy number variation in loci associated with cytochrome P450-mediated detoxification in a generalist arthropod pest. PLoS Genet. 2021;17: e1009422. doi:10.1371/journal.pgen.1009422

54. Villacis-Perez E, Snoeck S, Kurlovs AH, Clark RM, Breeuwer JAJ, Van Leeuwen T. Adaptive divergence and post-zygotic barriers to gene flow between sympatric populations of a herbivorous mite. Commun Biol. 2021;4: 853. doi:10.1038/s42003-021-02380-y

55. De Beer B, Villacis-Perez E, Khalighi M, Saalwaechter C, Vandenhole M, Jonckheere W, et al. QTL mapping suggests that both cytochrome P450-mediated detoxification and target-site resistance are involved in fenbutatin oxide resistance in *Tetranychus urticae*. Insect Biochem Mol Biol. 2022; 103757. doi:10.1016/j.ibmb.2022.103757

56. Bajda S, Dermauw W, Panteleri R, Sugimoto N, Douris V, Tirry L, et al. A mutation in the PSST homologue of complex I (NADH:ubiquinone oxidoreductase) from *Tetranychus urticae* is associated with resistance to METI acaricides. Insect Biochem Mol Biol. 2017;80: 79–90. doi:10.1016/j.ibmb.2016.11.010

57. Dermauw W, Ilias A, Riga M, Tsagkarakou A, Grbić M, Tirry L, et al. The cys-loop ligand-gated ion channel gene family of *Tetranychus urticae*: Implications for acaricide toxicology and a novel mutation associated with abamectin resistance. Insect Biochem Mol Biol. 2012;42: 455–465. doi:10.1016/j.ibmb.2012.03.002

58. Hubert J, Nesvorna M, Kamler M, Kopecky J, Tyl J, Titera D, et al. Point mutations in the sodium channel gene conferring tau-fluvalinate resistance in *Varroa destructor*: point mutations in *VmNa*. Pest Manag Sci. 2014;70: 889–894. doi:10.1002/ps.3679

59. González-Cabrera J, Rodríguez-Vargas S, Davies TGE, Field LM, Schmehl D, Ellis JD, et al. Novel mutations in the voltage-gated sodium channel of pyrethroid-resistant *Varroa destructor* populations from the Southeastern USA. PLoS One. 2016;11: e0155332. doi:10.1371/journal.pone.0155332

60. Vlogiannitis S, Jonckheere W, Laget D, de Graaf DC, Vontas J, Van Leeuwen T. Pyrethroid target-site resistance mutations in populations of the honey bee parasite *Varroa destructor* (Acari: Varroidae) from Flanders, Belgium. Exp Appl Acarol. 2021;85: 205–221. doi:10.1007/s10493-021-00665-9

61. Van Leeuwen T, Dermauw W, Mavridis K, Vontas J. Significance and interpretation of molecular diagnostics for insecticide resistance management of agricultural pests. Curr Opin Insect Sci. 2020;39: 69–76. doi:10.1016/j.cois.2020.03.006

62. Xue W, Lu X, Mavridis K, Vontas J, Jonckheere W, Van Leeuwen T. The H92R substitution in PSST is a reliable diagnostic biomarker for predicting resistance to mitochondrial electron transport inhibitors of complex I in European populations of *Tetranychus urticae*. Pest Manag Sci. 2022; ps.7007. doi:10.1002/ps.7007

63. Riga M, Bajda S, Themistokleous C, Papadaki S, Palzewicz M, Dermauw W, et al. The relative contribution of target-site mutations in complex acaricide resistant phenotypes as assessed by marker assisted backcrossing in *Tetranychus urticae*. Sci Rep. 2017;7. doi:10.1038/s41598-017-09054-y

64. Van Leeuwen T, Demaeght P, Osborne EJ, Dermauw W, Gohlke S, Nauen R, et al. Population bulk segregant mapping uncovers resistance mutations and the mode of action of a chitin synthesis inhibitor in arthropods. Proc Natl Acad Sci USA. 2012;109: 4407–4412. doi:10.1073/pnas.1200068109

65. Demaeght P, Osborne EJ, Odman-Naresh J, Grbić M, Nauen R, Merzendorfer H, et al. High resolution genetic mapping uncovers chitin synthase-1 as the target-site of the structurally diverse mite growth inhibitors clofentezine, hexythiazox and etoxazole in *Tetranychus urticae*. Insect Biochem Mol Biol. 2014;51: 52–61. doi:10.1016/j.ibmb.2014.05.004

66. Rioja C, Zhurov V, Bruinsma K, Grbic M, Grbic V. Plant-herbivore interactions: a case of an extreme generalist, the two-spotted spider mite *Tetranychus urticae*. Mol Plant Microbe Interact. 2017;30: 935–945. doi:10.1094/MPMI-07-17-0168-CR

67. Demaeght P, Dermauw W, Tsakireli D, Khajehali J, Nauen R, Tirry L, et al. Molecular analysis of resistance to acaricidal spirocyclic tetronic acids in *Tetranychus urticae*: CYP392E10 metabolizes spirodiclofen, but not its corresponding enol. Insect Biochem Mol Biol. 2013;43: 544–554. doi:10.1016/j.ibmb.2013.03.007

68. Riga M, Tsakireli D, Ilias A, Morou E, Myridakis A, Stephanou EG, et al. Abamectin is metabolized by CYP392A16, a cytochrome P450 associated with high levels of acaricide resistance in *Tetranychus urticae*. Insect Biochem Mol Biol. 2014;46: 43–53. doi:10.1016/j.ibmb.2014.01.006

69. Papapostolou KM, Riga M, Samantsidis G-R, Skoufa E, Balabanidou V, Van Leeuwen T, et al. Over-expression in *cis* of the midgut P450 CYP392A16 contributes to abamectin resistance in *Tetranychus urticae*. Insect Biochem Mol Biol. 2022;142: 103709. doi:10.1016/j.ibmb.2021.103709

70. Helle W, Pieterse AH. Genetic affinities between adjacent populations of spider mites (*Tetranychus urticae* Koch). Entomol Exp Appl. 1965;8: 305–308. doi:10.1111/j.1570-7458.1965.tb00866.x

71. Navajas M, Tsagkarakov A, Lagnel J, Perrot-Minnot M-J. Genetic differentiation in *Tetranychus urticae* (Acari: Tetranychidae): polymorphism, host races or sibling species? Exp Appl Acarol. 2000;24: 365–376. doi:10.1023/A:1006432604611

72. Cruz MA, Magalhães S, Sucena É, Zélé F. *Wolbachia* and host intrinsic reproductive barriers contribute additively to postmating isolation in spider mites. Evolution. 2021;75: 2085–2101. doi:10.1111/evo.14286

73. Wu TD, Nacu S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics. 2010;26: 873–881. doi:10.1093/bioinformatics/btq057

74. Wu TD, Reeder J, Lawrence M, Becker G, Brauer MJ. GMAP and GSNAP for genomic sequence alignment: enhancements to speed, accuracy, and functionality. In: Mathé E, Davis S, editors. Statistical Genomics. New York, NY: Springer New York; 2016. pp. 283–334. doi:10.1007/978-1-4939-3578-9_15

75. Fellous S, Angot G, Orsucci M, Migeon A, Auger P, Olivieri I, et al. Combining experimental evolution and field population assays to study the evolution of host range breadth. J Evol Biol. 2014;27: 911–919. doi:10.1111/jeb.12362

76. Springer NM, Stupar RM. Allele-specific expression patterns reveal biases and embryo-specific parent-of-origin effects in hybrid maize. Plant Cell. 2007;19: 2391–2402. doi:10.1105/tpc.107.052258

77. Villacis-Perez E, Alba JM, Cotte J, van Loon Z, Breeuwer JAJ, Van Leeuwen T. Interactions with plant defences isolate sympatric populations of an herbivorous mite. Front Ecol Evol. 2022;10: 819894. doi:10.3389/fevo.2022.819894

78. Gould BA, Chen Y, Lowry DB. Gene regulatory divergence between locally adapted ecotypes in their native habitats. Mol Ecol. 2018;27: 4174–4188. doi:10.1111/mec.14852

79. Cooper TF, Rozen DE, Lenski RE. Parallel changes in gene expression after 20,000 generations of evolution in *Escherichia coli*. Proc Natl Acad Sci USA. 2003;100: 1072–1077. doi:10.1073/pnas.0334340100

80. Stern DL, Orgogozo V. Is genetic evolution predictable? Science. 2009;323: 746–751. doi:10.1126/science.1158997

81. Smith LB, Tyagi R, Kasai S, Scott JG. CYP-mediated permethrin resistance in *Aedes aegypti* and evidence for *trans*-regulation. Foster GM, editor. PLoS Negl Trop Dis. 2018;12: e0006933. doi:10.1371/journal.pntd.0006933

82. Duveau F, Zande PV, Metzger BPH, Diaz CJ, Walker EA, Tryban S, et al. Mutational sources of *trans*- regulatory variation affecting gene expression in *Saccharomyces cerevisiae*. Elife. 2021. doi:10.7554/eLife.67806

83. King-Jones K, Horner MA, Lam G, Thummel CS. The DHR96 nuclear receptor regulates xenobiotic responses in *Drosophila*. Cell Metab. 2006;4: 37–48. doi:10.1016/j.cmet.2006.06.006

84. Pascussi J-M, Gerbal-Chaloin S, Duret C, Daujat-Chavanieu M, Vilarem M-J, Maurel P. The tangle of nuclear receptors that controls xenobiotic metabolism and transport: cross-talk and consequences. Annu Rev Pharmacol Toxicol. 2008;48: 1–32. doi:10.1146/annurev.pharmtox.47.120505.105349

85. Riga M, Myridakis A, Tsakireli D, Morou E, Stephanou EG, Nauen R, et al. Functional characterization of the *Tetranychus urticae* CYP392A11, a cytochrome P450 that hydroxylates the METI acaricides cyenopyrafen and fenpyroximate. Insect Biochem Mol Biol. 2015;65: 91–99. doi:10.1016/j.ibmb.2015.09.004

86. Wilding CS, Smith I, Lynd A, Yawson AE, Weetman D, Paine MJI, et al. A *cis*-regulatory sequence driving metabolic insecticide resistance in mosquitoes: functional characterisation and signatures of selection. Insect Biochem Mol Biol. 2012;42: 699–707. doi:10.1016/j.ibmb.2012.06.003

87. Pang R, Li Y, Dong Y, Liang Z, Zhang Y, Zhang W. Identification of promoter polymorphisms in the cytochrome P450 *CYP6AY1* linked with insecticide resistance in the brown planthopper, *Nilaparvata lugens*. Insect Mol Biol. 2014;23: 768–778. doi:10.1111/imb.12121

88. Daborn PJ, Yen JL, Bogwitz MR, Le Goff G, Feil E, Jeffers S, et al. A single P450 allele associated with insecticide resistance in *Drosophila*. Science. 2002;297: 2253–2256. doi:10.1126/science.1074170

89. Sayyed AH, Wright DJ. Cross-resistance and inheritance of resistance to *Bacillus thuringiensis* toxin Cry1Ac in diamondback moth (*Plutella xylostella* L) from lowland Malaysia. Pest Manag Sci. 2001;57: 413–421. doi:10.1002/ps.313

90. Wang YH, Liu XG, Zhu YC, Wu SG, Li SY, Chen WM, et al. Inheritance mode and realized heritability of resistance to imidacloprid in the brown planthopper, *Nilaparvata lugens* (Stål) (Homoptera: Delphacidae). Pest Manag Sci. 2009;65: 629–634. doi:10.1002/ps.1720

91. Diptaningsari D, Trisyono YA, Purwantoro A, Wijonarko A. Inheritance and realized heritability of resistance to imidacloprid in the brown planthopper, *Nilaparvata lugens* (Hemiptera: Delphacidae), from Indonesia. J Econ Entomol. 2019;112: 1831–1837. doi:10.1093/jee/toz090

92. Mansoor MM, Shad SA. Methoxyfenozide tolerance in *Chrysoperla carnea*: inheritance, dominance and preliminary detoxification mechanisms. PLoS One. 2022;17: e0265304. doi:10.1371/journal.pone.0265304

93. LeOra. POLO-Plus 1.0 Probit and Logit Analysis. Petaluma: LeOra Software; 2006.

94. Robertson JL, Savin NE, Savin NE, Preisler HK. Bioassays with arthropods. 2nd ed. CRC Press; 2007. doi:10.1201/9781420004045

95. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available: https://www.R-project.org/

96. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:13033997 [q-bio]. 2013. Available: http://arxiv.org/abs/1303.3997

97. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25: 2078–2079. doi:10.1093/bioinformatics/btp352

98. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20: 1297–1303. doi:10.1101/gr.107524.110

99. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43: 491–498. doi:10.1038/ng.806

100. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w ^{1118}□; iso-2; iso-3. Fly. 2012;6: 80–92. doi:10.4161/fly.19695

101. Sterck L, Billiau K, Abeel T, Rouzé P, Van de Peer Y. ORCAE: online resource for community annotation of eukaryotes. Nat Methods. 2012;9: 1041–1041. doi:10.1038/nmeth.2242

102. Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics. 2012;28: 3326–3328. doi:10.1093/bioinformatics/bts606

103. Zheng X, Gogarten SM, Lawrence M, Stilp A, Conomos MP, Weir BS, et al. SeqArray—a storage-efficient high-performance data format for WGS variant calls. Bioinformatics. 2017;33: 2251–2257. doi:10.1093/bioinformatics/btx145

104. Cowles CR, Hirschhorn JN, Altshuler D, Lander ES. Detection of regulatory variation in mouse genes. Nat Genet. 2002;32: 432–437. doi:10.1038/ng992

105. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31: 166–169. doi:10.1093/bioinformatics/btu638

106. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15: 1–21. doi:10.1186/s13059-014-0550-8

107. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics. 2009;25: 1422–1423. doi:10.1093/bioinformatics/btp163

108. Stephens M. False discovery rates: a new deal. Biostatistics. 2017;18: 275–294. doi:10.1093/biostatistics/kxw041

109. Stone BF. A formula for determining degree of dominance in cases of monofactorial inheritance of resistance to chemicals. Bull World Health Organ. 1968;38: 325–326.

110. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30: 1236. doi:10.1093/bioinformatics/btu031

111. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32: 1792–1797. doi:10.1093/nar/gkh340

112. Huerta-Cepas J, Dopazo J, Gabaldón T. ETE: a python environment for tree exploration. BMC Bioinformatics. 2010;11: 24. doi:10.1186/1471-2105-11-24

113. Hunter JD. Matplotlib: a 2D graphics environment. Comput Sci Eng. 2007;9: 90–95. doi:10.1109/MCSE.2007.55

Figures and legends

Fig 1. **Genetic and gene expression relatedness among five inbred *T. urticae* strains.** (A) A principal component (PC) analysis based on genetic variation (SNPs) shows relatedness among the five strains, as indicated by points with attending inbred strain names. (B) A PC analysis based on transcriptome profiles shows the similarities among the five strains and the F1 hybrids developed from crosses of four strains (ROS-ITi, JP-RRi, MAR-ABi, and MR-VPi) to a common parent (SOL-BEi). For F1 samples, labeling is based on the parent that varies (i.e., the non-SOL-BEi parent). Transcriptome data was collected from 4-7 day-old adult females, and each dot represents a biological replicate (4-5 replicates were used for each genotype). For both panels, PC1 and PC2 are indicated on the x- and y-axes, respectively (the percent of variance explained by each PC is given in parenthesis).

Fig 2. **Resistance phenotypes of inbred strains to twelve compounds and their performance on host plants.** (A) A heatmap showing resistance of five inbred *T. urticae* strains (top) to 12 acaricide compounds (left; IRAC pesticide classes are indicated at right). Log-transformed resistance ratios (RRs) were calculated based on the least resistant strain (scale, bottom). Within each cell, LC_{50} values (mg L^{-1}) are shown above the log-transformed RRs (in parentheses). For a given compound, strains for which LC_{50} values are not significantly different from each other have the same number of plus signs (indicated between LC_{50} values and the log-transformed RR values); the number of plus signs negatively correlates with resistance level (i.e., strains with a single plus sign have the highest LC_{50} values for a given acaricide, while strains with the highest number of plus signs have the lowest). By strain and compound, an asterisk (top right corner) denotes the presence of established target-site resistance mutations in *T. urticae* (S2 Table); the presence of target site mutations studied in other species, but not validated in *T. urticae*, are denoted by a question mark (bifenthrin only). All phenotypic data is from the current study, excepting that for fenbutatin oxide in strains SOL-BEi and MAR-ABi, which was reported previously by De Beer et al. [55]. (B) The total number of eggs laid by sets of ten mites (each a single replicate) on the

adaxial side of bean, cucumber, tomato, prune and potato leaf disks over a 3-day period is as indicated (legend, lower right). The data is displayed using boxplots with an overlay of data points (the number of replicates is given in parentheses). *P*-values (*p*) displayed are from an ANOVA; all significant pairwise differences among strains by host plant are indicated by asterisks ($p \leq 0.05$: *, $p \leq 0.01$: **, $p \leq 0.001$: ***, and $p \leq 0.0001$: ****, with post-hoc Tukey test to adjust for multiple testing).

Fig 3. Regulatory modes associated with gene expression variation in comparisons of parental strains and allele-specific expression in F1s. (A-D) For the comparisons of ROS-ITi (A), JP-RRi (B), MAR-ABi (C), and MR-VPi (D) to SOL-BEi, scatter plots show $\log_2$ fold change ($\log_2$FC) for genes with assigned genetic modes of control between the respective parents (y-axis) versus respective $\log_2$FC for allele-specific expression in F1s (x-axis) (for this analysis, genes assigned to conserved or ambiguous modes are not shown; $\log_2$FC values falling outside the plotting scale of -10 to 10 were rescaled to these limits for plotting). Each point represents a single gene, with color coding by *cis*, *trans*, synergism, antagonism, and compensatory mode of genetic control as indicated (see insets for color codes and respective gene numbers; *n*, total sample size, upper left). (E-H) Respective plots for panels A-D except that only genes in the detoxification/host/peptidase (D+H+P) gene set are plotted, and with selection for genes with moderate to large absolute value of $\log_2$FC (> 1) along one or more axes. Different symbols and outline colors denote membership in the various D+H+P gene families or gene set classifications (see legends to the right of each panel). The interiors of plotting symbols are color coded by the regulatory categories as indicated in the insets (*n*, total sample size, upper left). For each gene family or for host genes, in the legends the three numbers in parentheses represent (*i*) the number of genes shown in each panel with *cis*, *trans*, synergism, antagonism, and compensatory mode assignments for which $|\log_2\text{FC}| > 1$, (*ii*) genes with the same selection criteria except with no $\log_2$FC cutoff applied, out of (*iii*) the total number of genes for which any mode of genetic control could be assigned (including conserved and ambiguous; see also S2 Fig). The data upon which this figure is based are provided in S4 Table.

Fig 4. Composition of genetic control categories stratified by magnitude of parental expression fold change differences. (A) For all genes for which *cis*, *trans*, synergism, antagonism or compensatory modes of expression control (legend, top right) could be assigned in pairwise strain comparisons, stacked barplots show respective percentages stratified by increasing absolute value of $\log_2$ fold change ($|\log_2 FC|$) bins for parental expression differences (denoted beneath each subpanel; the four genotypic comparisons, as well as respective sample sizes, *n*, are indicated above each subpanel). (B) Same as for panel (A) except for the subset of genes in the detoxification/host/peptidase (D+H+P) gene set. The data upon which this figure is based are provided in S4 Table.

Fig 5. The contribution of *cis* and *trans* effects for CYP gene expression variation among pairwise strain comparisons. A heatmap (center right; scale, bottom left) shows $\log_2$ fold change ($\log_2 FC$) values for CYPs between parents (a combination of *cis* and *trans* effects, denoted "*cis* + *trans*"), for ASE in F1s (which correspond to *cis* effects only, denoted *cis*), and for the *trans* contribution as inferred from the *cis* + *trans* and *cis* $\log_2 FC$ values (see Methods). These groupings are as indicated at the top, and the pairs of strains for the four comparisons are indicated at bottom (exact $\log_2 FC$ values are given internal to each cell, and if significant, an asterisk is used; see also S4 Table). Genes with $\log_2 FC > 1$ attributed to *trans* in at least one of the comparisons are in bold, and genes with $\log_2 FC < -1$ attributed to *trans* in at least one of the comparisons are underlined. CYP gene names are given to the left, and are ordered based on a neighbor-joining phylogenetic tree (far left). CYP clan membership is indicated at the far right, with branches in the tree also color-coded respectively. For inclusion of a CYP in the analysis, a genetic mode of control had to be assigned in at least one comparison. Where information for a gene was not available in a specific genotypic comparison (see Methods), cells are colored in gray.

Fig 6. **Inheritance of gene expression and its association with regulatory categories for gene expression variation.** (A-B) Stacked barplots display the composition of each inheritance mode for comparisons of four strains (ROS-ITi, JP-RRi, MAR-ABi, and MR-VPi, indicated as P1, see labels at

bottom) to a common parent (SOL-BEi, indicated as P2) for all genes (A) and for those in the detoxification/host/peptidase (D+H+P) gene set (B). The following inheritance classifications were made: (*i*) additive, where that of F1 equals the average expression of the respective parents, (*ii*) dominant, where that of the F1 equals expression of one parent (the direction of dominance is as indicated in the legend at the far left), and (*iii*) transgressive, for which F1 values are outside the range of the parents (the direction of transgression is defined as either up or down with respect to the parental range). The number of genes (y-axis) for P1>P2 (above 0) and P1<P2 (below 0) are as indicated above or below the bars in the plot, with coloring of stacked bars denoting inheritance classifications (see legend, far right; within the stacked bar plots, exact percentages are given for modes > 5%). In panels A and B, transgressive inheritance for which P1=P2 (no expression difference between parents) is not indicated. (C) Line plot displaying the relative proportions of additive and dominant inheritance modes (y-axis) stratified by absolute values of $\log_2$ fold change in comparisons between parents ($|\log_2(P1/P2)|$, x-axis). The different line types represent the four comparisons, and different colors denote additive or dominant modes (see legend at right). (D) The association between each combination of mode of genetic control (*cis*, *trans*, synergism, and antagonism) and inheritance mode (additive, dominant, and transgressive). The significance of enrichment (blue) or depletion (red) is denoted using asterisks (Chi-square test of independence followed by Benjamini-Hochberg multiple test corrections; adjusted $p < 0.001$: ***; adjusted $p < 0.01$: **; adjusted $p < 0.05$: *). The data upon which this figure is based are provided in S7 Table.

Supporting information

S1 Fig. Gene numbers for inclusion in assessing genetic mode of control of expression variation. (A) A Venn diagram shows the number of intact+single copy genes available for potential assessment of genetic mode of control by comparison of strains ROS-ITi, JP-RRi, MAR-ABi, and MR-VPi to a single parent strain (SOL-BEi), and the overlap among the comparisons (for details, see Methods, and S3 Table). (B) A Venn diagram after panel (A) but indicating the number of genes in each comparison, and the overlap among comparisons, where an assignment of genetic mode of control could be made (categories *cis*, *trans*, synergism, antagonism, compensatory, conserved, and ambiguous). See S4 Table for the respective genes by comparison.

S2 Fig. The composition of genetic mode of gene expression control in pairwise strain comparisons by gene group. For pairwise comparisons of each of ROS-ITi (A), JP-RRi (B), MAR-ABi (C), and MR-VPi (D) to SOL-BEi, the number of genes assigned into one of the seven genetic control categories is indicated above the stacked barplots (categories *cis*, *trans*, synergism, antagonism, compensatory, conserved, and ambiguous are as indicated by color in the stacks, see the legend at top). In each of panels A-D, classifications are shown for all genes in the left subpanel, and at right for genes in the detoxification/host/peptidase (D+H+P) gene set sorted by gene family or group (as in the case of host genes). The data upon which this figure is based are provided in S4 Table.

S3 Fig. The contribution of *cis* and *trans* effects for dioxygenase gene expression variation among pairwise strain comparisons. A heatmap (center right; scale, bottom left) shows $\log_2$ fold change ($\log_2$FC) values for dioxygenase (DOG) genes between parents (a combination of *cis* and *trans* effects, denoted "*cis* + *trans*"), for ASE in F1s (which correspond to *cis* effects only, denoted *cis*), and for the *trans* contribution as inferred from the *cis* + *trans* and *cis* $\log_2$FC values (see Methods). These groupings are as indicated at the top, and the pairs of strains for the four comparisons are indicated at bottom (exact $\log_2$FC values are given internal to each cell, and if significant, an asterisk is used; see also S4 Table).

Genes are ordered based on a neighbor-joining phylogenetic tree as indicated on the left. For inclusion of a DOG in the analysis, a genetic mode of control had to have been assigned in at least one comparison. Where information for a gene was not available in a specific genotypic comparison, cells are colored in gray.

S4 Fig. The contribution of *cis* and *trans* effects for lipocalin gene expression variation among pairwise strain comparisons. A heatmap (center right; scale, bottom left) shows $\log_2$ fold change ($\log_2$FC) values for lipocalin genes between parents (a combination of *cis* and *trans* effects, denoted "*cis* + *trans*"), for ASE in F1s (which correspond to *cis* effects only, denoted *cis*), and for the *trans* contribution as inferred from the *cis* + *trans* and *cis* $\log_2$FC values (see Methods). These groupings are as indicated at the top, and the pairs of strains for the four comparisons are indicated at bottom (exact $\log_2$FC values are given internal to each cell, and if significant, an asterisk is used; see also S4 Table). Genes are ordered based on a neighbor-joining phylogenetic tree as indicated on the left. For inclusion of a lipocalin in the analysis, a genetic mode of control had to be assigned in at least one comparison. Where information for a gene was not available in a specific genotypic comparison, cells are colored in gray.

S5 Fig. The contribution of *cis* and *trans* effects for short-chain dehydrogenase/reductase gene expression variation among pairwise strain comparisons. A heatmap (center right; scale, bottom left) shows $\log_2$ fold change ($\log_2$FC) values for short-chain dehydrogenases/reductase (SDR) genes between parents (a combination of *cis* and *trans* effects, denoted "*cis* + *trans*"), for ASE in F1s (which correspond to *cis* effects only, denoted *cis*), and for the *trans* contribution as inferred from the *cis* + *trans* and *cis* $\log_2$FC values (see Methods). These groupings are as indicated at the top, and the pairs of strains for the four comparisons are indicated at bottom (exact $\log_2$FC values are given internal to each cell, and if significant, an asterisk is used; see also S4 Table). Genes are ordered based on a neighbor-joining phylogenetic tree as indicated on the left. For inclusion of a SDR in the analysis, a genetic mode of control had to

be assigned in at least one comparison. Where information for a gene was not available in a specific genotypic comparison, cells are colored in gray.

S1 Table. Chemical compounds used in the bioassays with concentration, corresponding IRAC group, mode of action (MoA) and time of mortality assessment

S2 Table. Target site mutations present by strain.

S3 Table. Putative intact+single copy genes in pairwise comparisons

S4 Table. Regulatory modes for gene expression variation in pairwise comparisons.

S5 Table. Genes in the detoxification/host/peptidase (D+H+P) gene set.

S6 Table. Enriched gene ontology (GO) terms for genes with extremes of *trans* genetic control.

S7 Table. Gene expression inheritance inferred in pairwise comparisons.

Fig 1

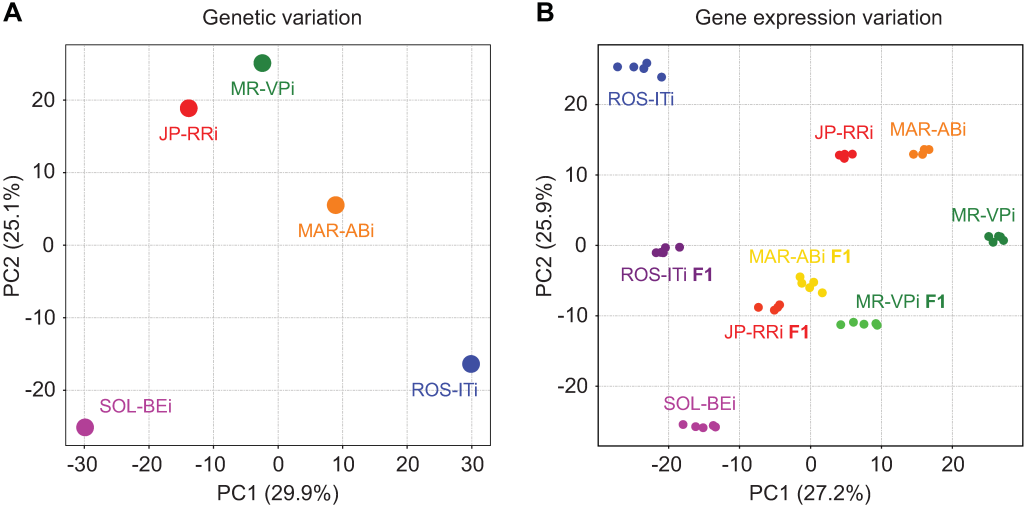

A
Genetic variation
MR-VPi
JP-RRi
MAR-ABi
ROS-ITi
SOL-BEi
PC2 (25.1%)
PC1 (29.9%)
20
10
0
-10
-20
-30
-20
-10
0
10
20
30
B
Gene expression variation
ROS-ITi
JP-RRi
MAR-ABi
MR-VPi
MAR-ABi F1
ROS-ITi F1
MR-VPi F1
JP-RRi F1
SOL-BEi
PC2 (25.9%)
PC1 (27.2%)
20
10
0
-10
-20
-20
-10
0
10
20

Fig 2

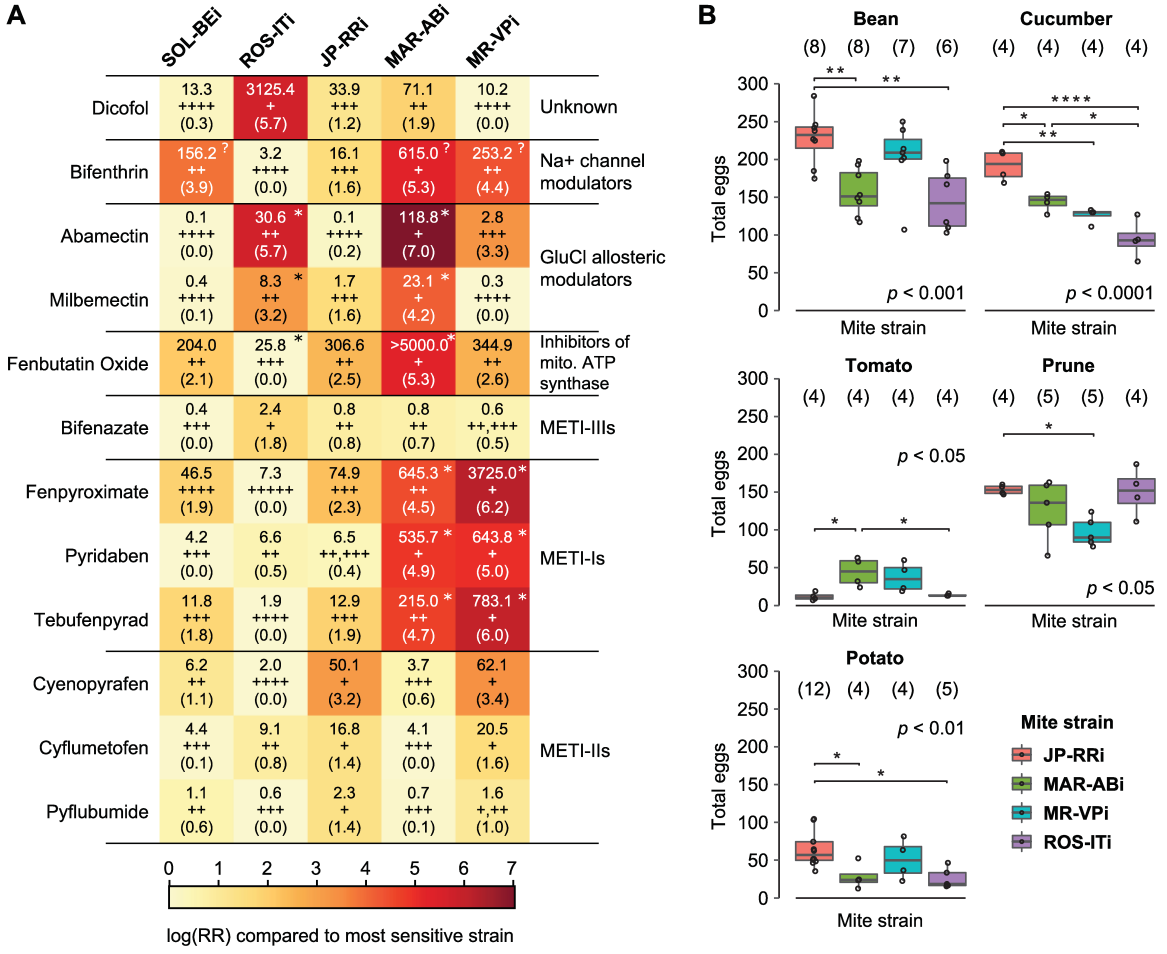

A
SOL-BEi
ROS-ITi
JP-RRi
MAR-ABi
MR-VPi
Dicofol
13.3 ++++ (0.3)
3125.4 + (5.7)
33.9 +++ (1.2)
71.1 ++ (1.9)
10.2 ++++ (0.0)
Unknown
Bifenthrin
156.2 ? ++ (3.9)
3.2 ++++ (0.0)
16.1 +++ (1.6)
615.0 ? + (5.3)
253.2 ? ++ (4.4)
Na+ channel modulators
Abamectin
0.1 ++++ (0.0)
30.6 * ++ (5.7)
0.1 ++++ (0.2)
118.8 * + (7.0)
2.8 +++ (3.3)
Milbemectin
0.4 ++++ (0.1)
8.3 * ++ (3.2)
1.7 +++ (1.6)
23.1 * + (4.2)
0.3 ++++ (0.0)
GluCl allosteric modulators
Fenbutatin Oxide
204.0 ++ (2.1)
25.8 * +++ (0.0)
306.6 ++ (2.5)
>5000.0 * + (5.3)
344.9 ++ (2.6)
Inhibitors of mito. ATP synthase
Bifenazate
0.4 +++ (0.0)
2.4 + (1.8)
0.8 ++ (0.8)
0.8 ++ (0.7)
0.6 ++,+++ (0.5)
METI-IIIs
Fenpyroximate
46.5 ++++ (1.9)
7.3 +++++ (0.0)
74.9 +++ (2.3)
645.3 * ++ (4.5)
3725.0 * + (6.2)
Pyridaben
4.2 +++ (0.0)
6.6 ++ (0.5)
6.5 ++,+++ (0.4)
535.7 * + (4.9)
643.8 * + (5.0)
Tebufenpyrad
11.8 +++ (1.8)
1.9 ++++ (0.0)
12.9 +++ (1.9)
215.0 * ++ (4.7)
783.1 * + (6.0)
METI-Is
Cyenopyrafen
6.2 ++ (1.1)
2.0 ++++ (0.0)
50.1 + (3.2)
3.7 +++ (0.6)
62.1 + (3.4)
Cyflumetofen
4.4 +++ (0.1)
9.1 ++ (0.8)
16.8 + (1.4)
4.1 +++ (0.0)
20.5 + (1.6)
Pyflubumide
1.1 ++ (0.6)
0.6 +++ (0.0)
2.3 + (1.4)
0.7 +++ (0.1)
1.6 +,++ (1.0)
METI-IIs
0 1 2 3 4 5 6 7
log(RR) compared to most sensitive strain
B
Bean
(8) (8) (7) (6)
Cucumber
(4) (4) (4) (4)
Total eggs
Mite strain
p < 0.001
p < 0.0001
Tomato
(4) (4) (4) (4)
Prune
(4) (5) (5) (4)
p < 0.05
Potato
(12) (4) (4) (5)
p < 0.01
Mite strain
JP-RRi
MAR-ABi
MR-VPi
ROS-ITi

Fig 3

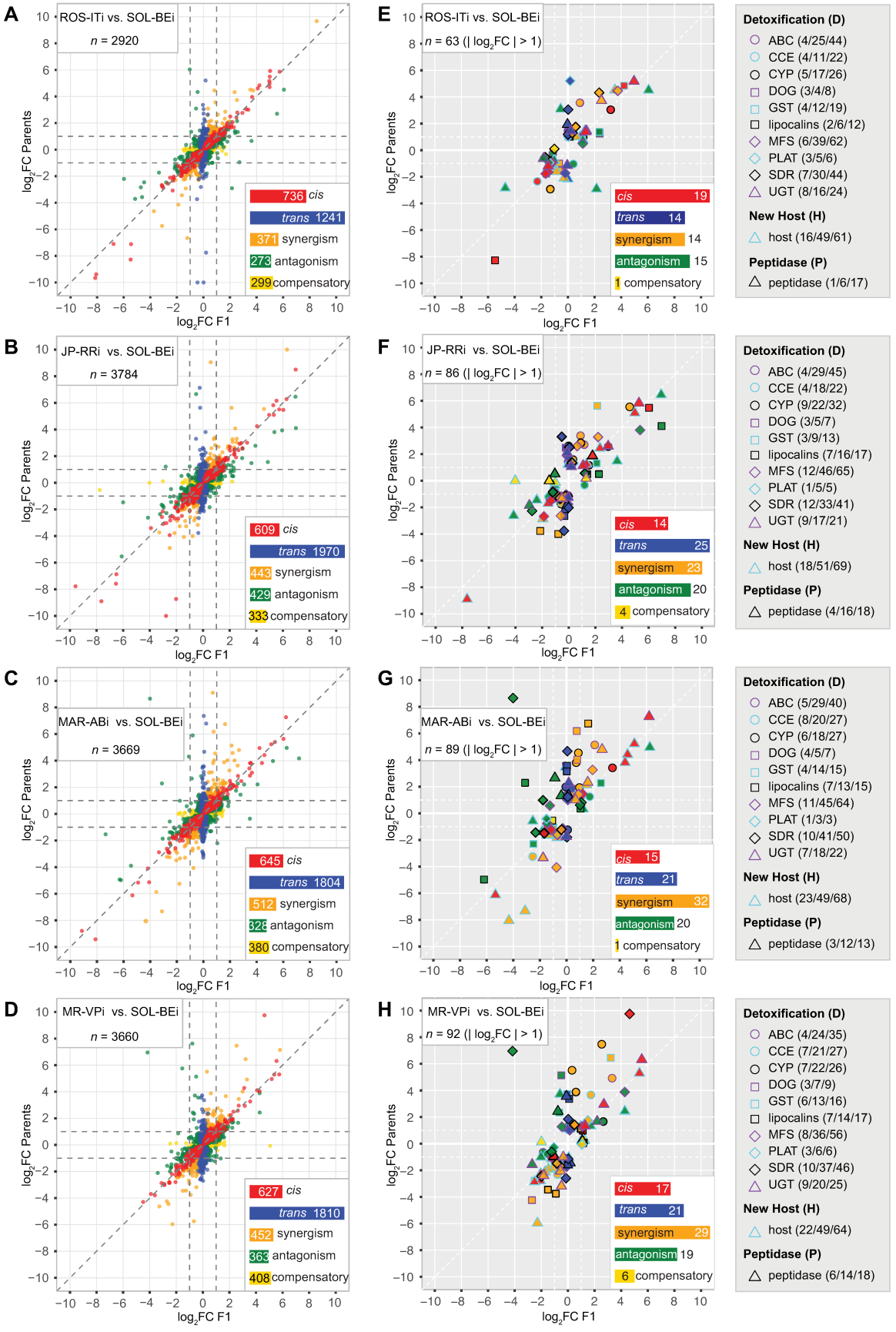

A
ROS-ITi vs. SOL-BEi
n = 2920
736 cis
trans 1241
371 synergism
273 antagonism
299 compensatory
B
JP-RRi vs. SOL-BEi
n = 3784
609 cis
trans 1970
443 synergism
429 antagonism
333 compensatory
C
MAR-ABi vs. SOL-BEi
n = 3669
645 cis
trans 1804
512 synergism
328 antagonism
380 compensatory
D
MR-VPi vs. SOL-BEi
n = 3660
627 cis
trans 1810
452 synergism
363 antagonism
408 compensatory
E
ROS-ITi vs. SOL-BEi
n = 63 (| log2FC | > 1)
cis 19
trans 14
synergism 14
antagonism 15
1 compensatory
F
JP-RRi vs. SOL-BEi
n = 86 (| log2FC | > 1)
cis 14
trans 25
synergism 23
antagonism 20
4 compensatory
G
MAR-ABi vs. SOL-BEi
n = 89 (| log2FC | > 1)
cis 15
trans 21
synergism 32
antagonism 20
1 compensatory
H
MR-VPi vs. SOL-BEi
n = 92 (| log2FC | > 1)
cis 17
trans 21
synergism 29
antagonism 19
6 compensatory
log2FC Parents
log2FC F1
Detoxification (D)
ABC (4/25/44)
CCE (4/11/22)
CYP (5/17/26)
DOG (3/4/8)
GST (4/12/19)
lipocalins (2/6/12)
MFS (6/39/62)
PLAT (3/5/6)
SDR (7/30/44)
UGT (8/16/24)
New Host (H)
host (16/49/61)
Peptidase (P)
peptidase (1/6/17)
Detoxification (D)
ABC (4/29/45)
CCE (4/18/22)
CYP (9/22/32)
DOG (3/5/7)
GST (3/9/13)
lipocalins (7/16/17)
MFS (12/46/65)
PLAT (1/5/5)
SDR (12/33/41)
UGT (9/17/21)
New Host (H)
host (18/51/69)
Peptidase (P)
peptidase (4/16/18)
Detoxification (D)
ABC (5/29/40)
CCE (8/20/27)
CYP (6/18/27)
DOG (4/5/7)
GST (4/14/15)
lipocalins (7/13/15)
MFS (11/45/64)
PLAT (1/3/3)
SDR (10/41/50)
UGT (7/18/22)
New Host (H)
host (23/49/68)
Peptidase (P)
peptidase (3/12/13)
Detoxification (D)
ABC (4/24/35)
CCE (7/21/27)
CYP (7/22/26)
DOG (3/7/9)
GST (6/13/16)
lipocalins (7/14/17)
MFS (8/36/56)
PLAT (3/6/6)
SDR (10/37/46)
UGT (9/20/25)
New Host (H)
host (22/49/64)
Peptidase (P)
peptidase (6/14/18)

Fig 4

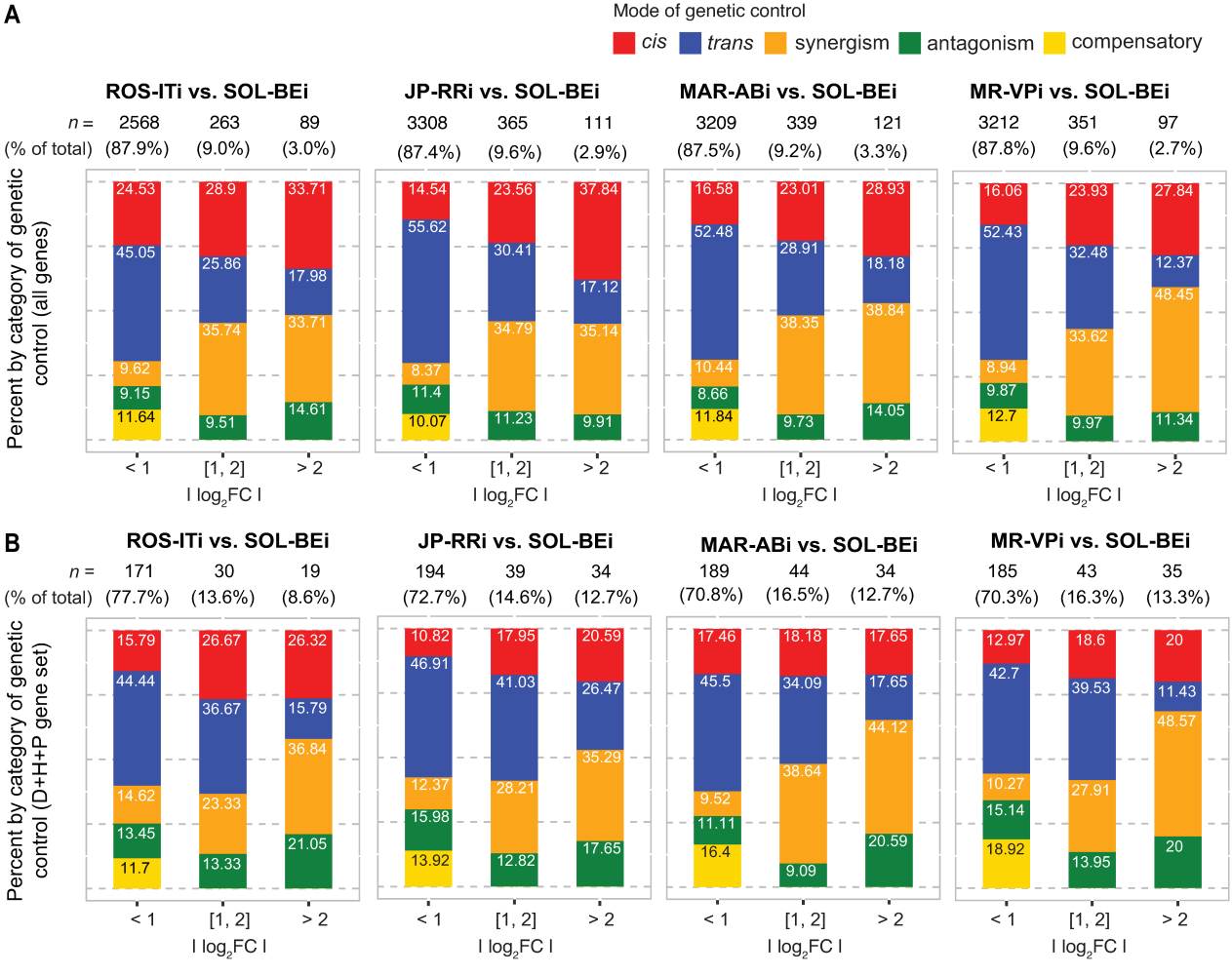

A
Mode of genetic control
cis
trans
synergism
antagonism
compensatory
ROS-ITi vs. SOL-BEi
JP-RRi vs. SOL-BEi
MAR-ABi vs. SOL-BEi
MR-VPi vs. SOL-BEi
n = 2568 263 89
(% of total) (87.9%) (9.0%) (3.0%)
3308 365 111
(87.4%) (9.6%) (2.9%)
3209 339 121
(87.5%) (9.2%) (3.3%)
3212 351 97
(87.8%) (9.6%) (2.7%)
Percent by category of genetic control (all genes)
24.53 45.05 9.62 9.15 11.64
28.9 25.86 35.74 9.51
33.71 17.98 33.71 14.61
14.54 55.62 8.37 11.4 10.07
23.56 30.41 34.79 11.23
37.84 17.12 35.14 9.91
16.58 52.48 10.44 8.66 11.84
23.01 28.91 38.35 9.73
28.93 18.18 38.84 14.05
16.06 52.43 8.94 9.87 12.7
23.93 32.48 33.62 9.97
27.84 12.37 48.45 11.34
< 1
[1, 2]
> 2
| log2FC |
B
ROS-ITi vs. SOL-BEi
JP-RRi vs. SOL-BEi
MAR-ABi vs. SOL-BEi
MR-VPi vs. SOL-BEi
n = 171 30 19
(% of total) (77.7%) (13.6%) (8.6%)
194 39 34
(72.7%) (14.6%) (12.7%)
189 44 34
(70.8%) (16.5%) (12.7%)
185 43 35
(70.3%) (16.3%) (13.3%)
Percent by category of genetic control (D+H+P gene set)
15.79 44.44 14.62 13.45 11.7
26.67 36.67 23.33 13.33
26.32 15.79 36.84 21.05
10.82 46.91 12.37 15.98 13.92
17.95 41.03 28.21 12.82
20.59 26.47 35.29 17.65
17.46 45.5 9.52 11.11 16.4
18.18 34.09 38.64 9.09
17.65 17.65 44.12 20.59
12.97 42.7 10.27 15.14 18.92
18.6 39.53 27.91 13.95
20 11.43 48.57 20
< 1
[1, 2]
> 2
| log2FC |

Fig 5

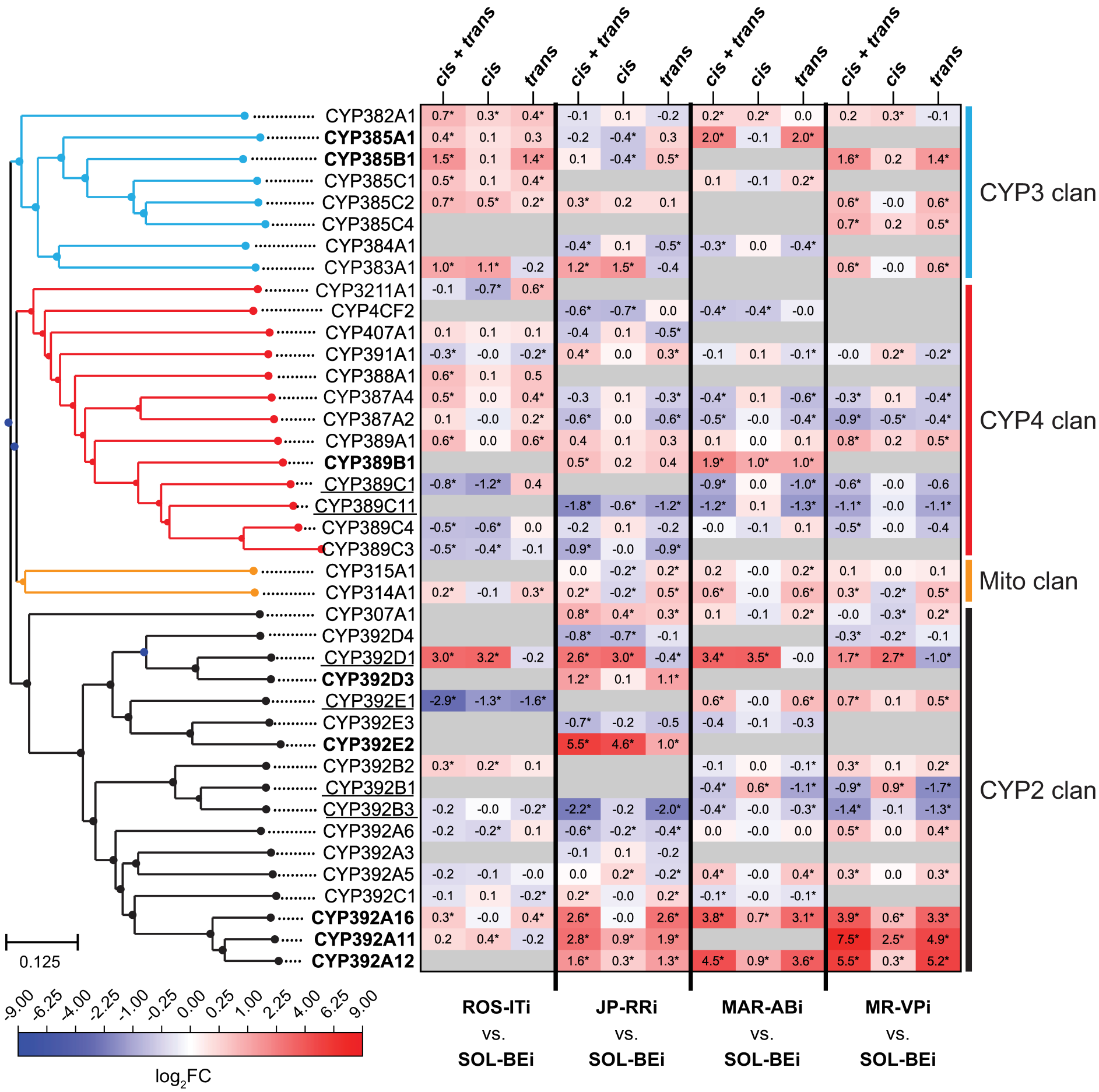
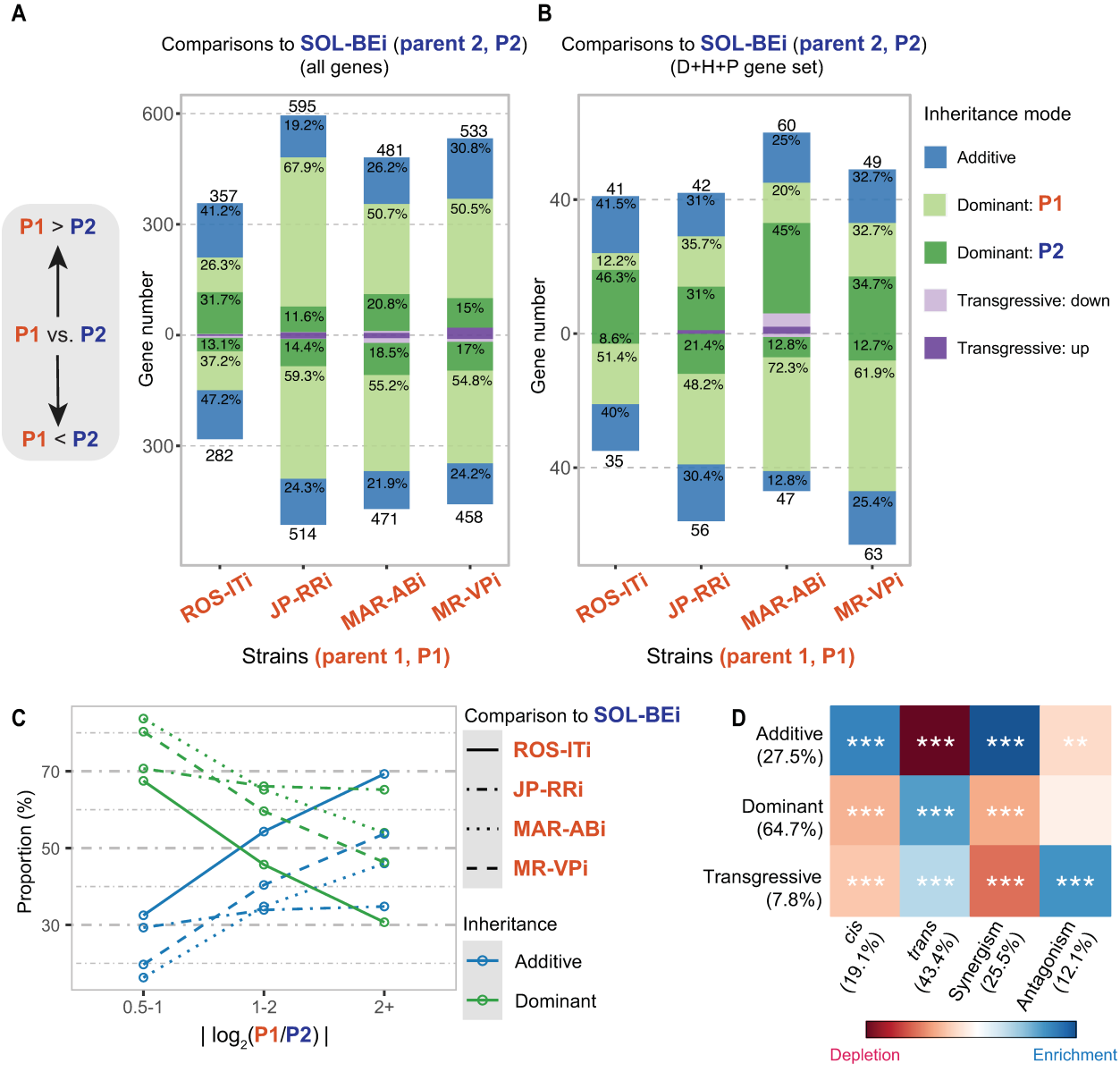
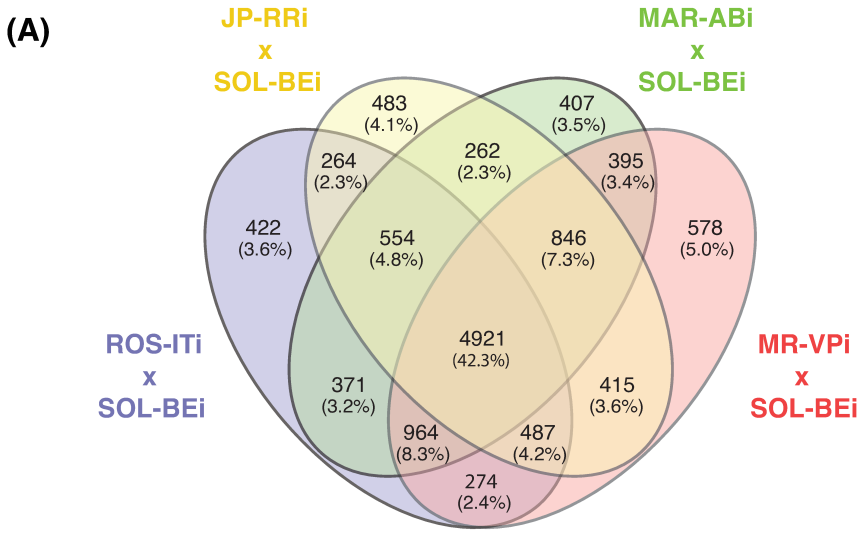
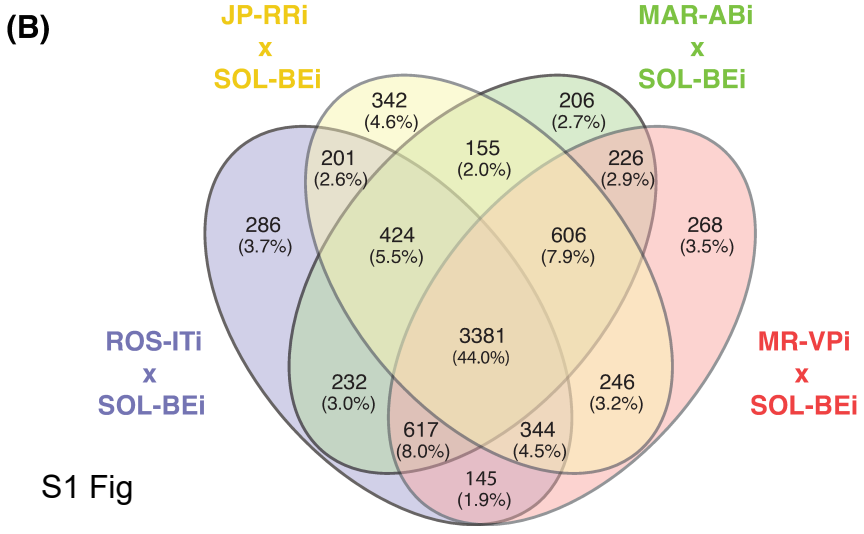

cis + trans
cis
trans
CYP382A1 0.7* 0.3* 0.4* -0.1 0.1 -0.2 0.2* 0.2* 0.0 0.2 0.3* -0.1
CYP385A1 0.4* 0.1 0.3 -0.2 -0.4* 0.3 2.0* -0.1 2.0*
CYP385B1 1.5* 0.1 1.4* 0.1 -0.4* 0.5* 1.6* 0.2 1.4*
CYP385C1 0.5* 0.1 0.4* 0.1 -0.1 0.2*
CYP385C2 0.7* 0.5* 0.2* 0.3* 0.2 0.1 0.6* -0.0 0.6*
CYP385C4 0.7* 0.2 0.5*
CYP384A1 -0.4* 0.1 -0.5* -0.3* 0.0 -0.4*
CYP383A1 1.0* 1.1* -0.2 1.2* 1.5* -0.4 0.6* -0.0 0.6*
CYP3211A1 -0.1 -0.7* 0.6*
CYP4CF2 -0.6* -0.7* 0.0 -0.4* -0.4* -0.0
CYP407A1 0.1 0.1 0.1 -0.4 0.1 -0.5*
CYP391A1 -0.3* -0.0 -0.2* 0.4* 0.0 0.3* -0.1 0.1 -0.1* -0.0 0.2* -0.2*
CYP388A1 0.6* 0.1 0.5
CYP387A4 0.5* 0.0 0.4* -0.3 0.1 -0.3* -0.4* 0.1 -0.6* -0.3* 0.1 -0.4*
CYP387A2 0.1 -0.0 0.2* -0.6* 0.0 -0.6* -0.5* -0.0 -0.4* -0.9* -0.5* -0.4*
CYP389A1 0.6* 0.0 0.6* 0.4 0.1 0.3 0.1 0.0 0.1 0.8* 0.2 0.5*
CYP389B1 0.5* 0.2 0.4 1.9* 1.0* 1.0*
CYP389C1 -0.8* -1.2* 0.4 -0.9* 0.0 -1.0* -0.6* -0.0 -0.6
CYP389C11 -1.8* -0.6* -1.2* -1.2* 0.1 -1.3* -1.1* -0.0 -1.1*
CYP389C4 -0.5* -0.6* 0.0 -0.2 0.1 -0.2 -0.0 -0.1 0.1 -0.5* -0.0 -0.4
CYP389C3 -0.5* -0.4* -0.1 -0.9* -0.0 -0.9*
CYP315A1 0.0 -0.2* 0.2* 0.2 -0.0 0.2* 0.1 0.0 0.1
CYP314A1 0.2* -0.1 0.3* 0.2* -0.2* 0.5* 0.6* -0.0 0.6* 0.3* -0.2* 0.5*
CYP307A1 0.8* 0.4* 0.3* 0.1 -0.1 0.2* -0.0 -0.3* 0.2*
CYP392D4 -0.8* -0.7* -0.1 -0.3* -0.2* -0.1
CYP392D1 3.0* 3.2* -0.2 2.6* 3.0* -0.4* 3.4* 3.5* -0.0 1.7* 2.7* -1.0*
CYP392D3 1.2* 0.1 1.1*
CYP392E1 -2.9* -1.3* -1.6* 0.6* -0.0 0.6* 0.7* 0.1 0.5*
CYP392E3 -0.7* -0.2 -0.5 -0.4 -0.1 -0.3
CYP392E2 5.5* 4.6* 1.0*
CYP392B2 0.3* 0.2* 0.1 -0.1 0.0 -0.1* 0.3* 0.1 0.2*
CYP392B1 -0.4* 0.6* -1.1* -0.9* 0.9* -1.7*
CYP392B3 -0.2 -0.0 -0.2* -2.2* -0.2 -2.0* -0.4* -0.0 -0.3* -1.4* -0.1 -1.3*
CYP392A6 -0.2 -0.2* 0.1 -0.6* -0.2* -0.4* 0.0 -0.0 0.0 0.5* 0.0 0.4*
CYP392A3 -0.1 0.1 -0.2
CYP392A5 -0.2 -0.1 -0.0 0.0 0.2* -0.2* 0.4* -0.0 0.4* 0.3* 0.0 0.3*
CYP392C1 -0.1 0.1 -0.2* 0.2* -0.0 0.2* -0.1* -0.0 -0.1*
CYP392A16 0.3* -0.0 0.4* 2.6* -0.0 2.6* 3.8* 0.7* 3.1* 3.9* 0.6* 3.3*
CYP392A11 0.2 0.4* -0.2 2.8* 0.9* 1.9* 7.5* 2.5* 4.9*
CYP392A12 1.6* 0.3* 1.3* 4.5* 0.9* 3.6* 5.5* 0.3* 5.2*
CYP3 clan
CYP4 clan
Mito clan
CYP2 clan
0.125
-9.00 -6.25 -4.00 -2.25 -1.00 -0.25 0.00 0.25 1.00 2.25 4.00 6.25 9.00
log2FC
ROS-ITi vs. SOL-BEi
JP-RRi vs. SOL-BEi
MAR-ABi vs. SOL-BEi
MR-VPi vs. SOL-BEi

Fig 6

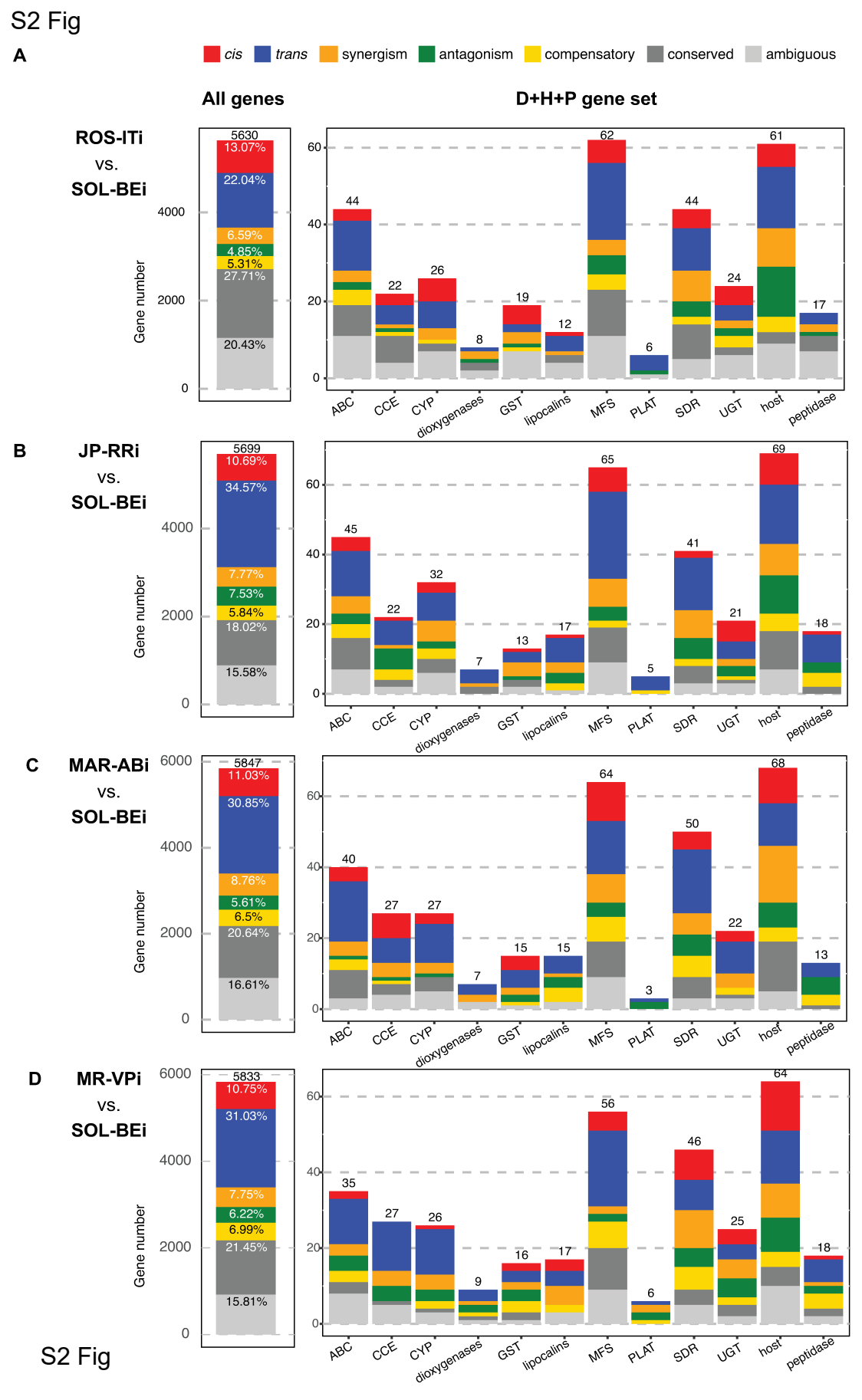

A
Comparisons to SOL-BEi (parent 2, P2)
(all genes)
P1 > P2
P1 vs. P2
P1 < P2
Gene number
600
300
0
300
357
41.2%
26.3%
31.7%
13.1%
37.2%
47.2%
282
595
19.2%
67.9%
11.6%
14.4%
59.3%
24.3%
514
481
26.2%
50.7%
20.8%
18.5%
55.2%
21.9%
471
533
30.8%
50.5%
15%
17%
54.8%
24.2%
458
ROS-ITi
JP-RRi
MAR-ABi
MR-VPi
Strains (parent 1, P1)
B
Comparisons to SOL-BEi (parent 2, P2)
(D+H+P gene set)
Gene number
40
0
40
41
41.5%
12.2%
46.3%
8.6%
51.4%
40%
35
42
31%
35.7%
31%
21.4%
48.2%
30.4%
56
60
25%
20%
45%
12.8%
72.3%
12.8%
47
49
32.7%
32.7%
34.7%
12.7%
61.9%
25.4%
63
ROS-ITi
JP-RRi
MAR-ABi
MR-VPi
Strains (parent 1, P1)
Inheritance mode
Additive
Dominant: P1
Dominant: P2
Transgressive: down
Transgressive: up
C
Proportion (%)
70
50
30
0.5-1
1-2
2+
| log2(P1/P2) |
Comparison to SOL-BEi
ROS-ITi
JP-RRi
MAR-ABi
MR-VPi
Inheritance
Additive
Dominant
D
Additive (27.5%)
Dominant (64.7%)
Transgressive (7.8%)
***
**
cis (19.1%)
trans (43.4%)
Synergism (25.5%)
Antagonism (12.1%)
Depletion
Enrichment

S1 Fig

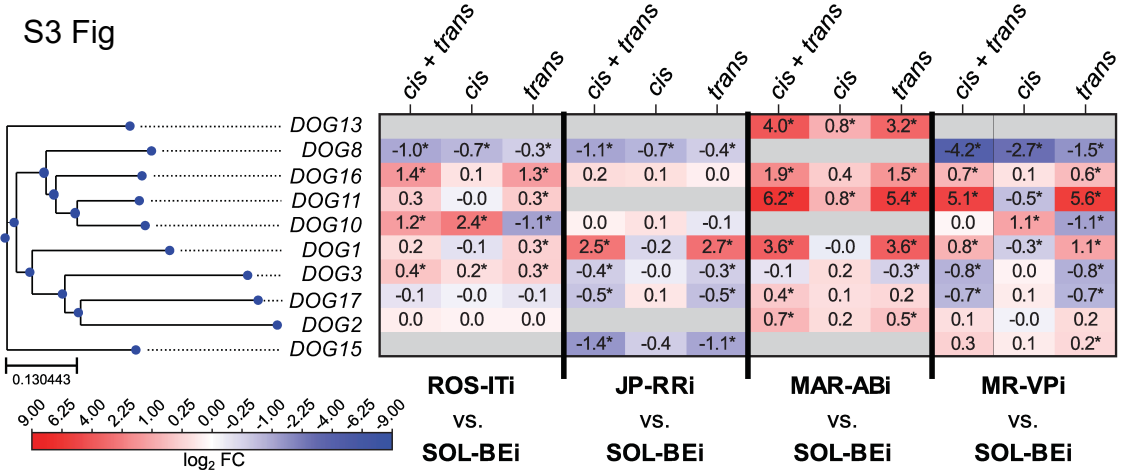

(A)
JP-RRi
x
SOL-BEi
MAR-ABi
x
SOL-BEi
ROS-ITi
x
SOL-BEi
MR-VPi
x
SOL-BEi
483
(4.1%)
407
(3.5%)
264
(2.3%)
262
(2.3%)
395
(3.4%)
422
(3.6%)
554
(4.8%)
846
(7.3%)
578
(5.0%)
4921
(42.3%)
371
(3.2%)
415
(3.6%)
964
(8.3%)
487
(4.2%)
274
(2.4%)


(B)
JP-RRi
x
SOL-BEi
MAR-ABi
x
SOL-BEi
ROS-ITi
x
SOL-BEi
MR-VPi
x
SOL-BEi
342
(4.6%)
206
(2.7%)
201
(2.6%)
155
(2.0%)
226
(2.9%)
286
(3.7%)
424
(5.5%)
606
(7.9%)
268
(3.5%)
3381
(44.0%)
232
(3.0%)
246
(3.2%)
617
(8.0%)
344
(4.5%)
145
(1.9%)

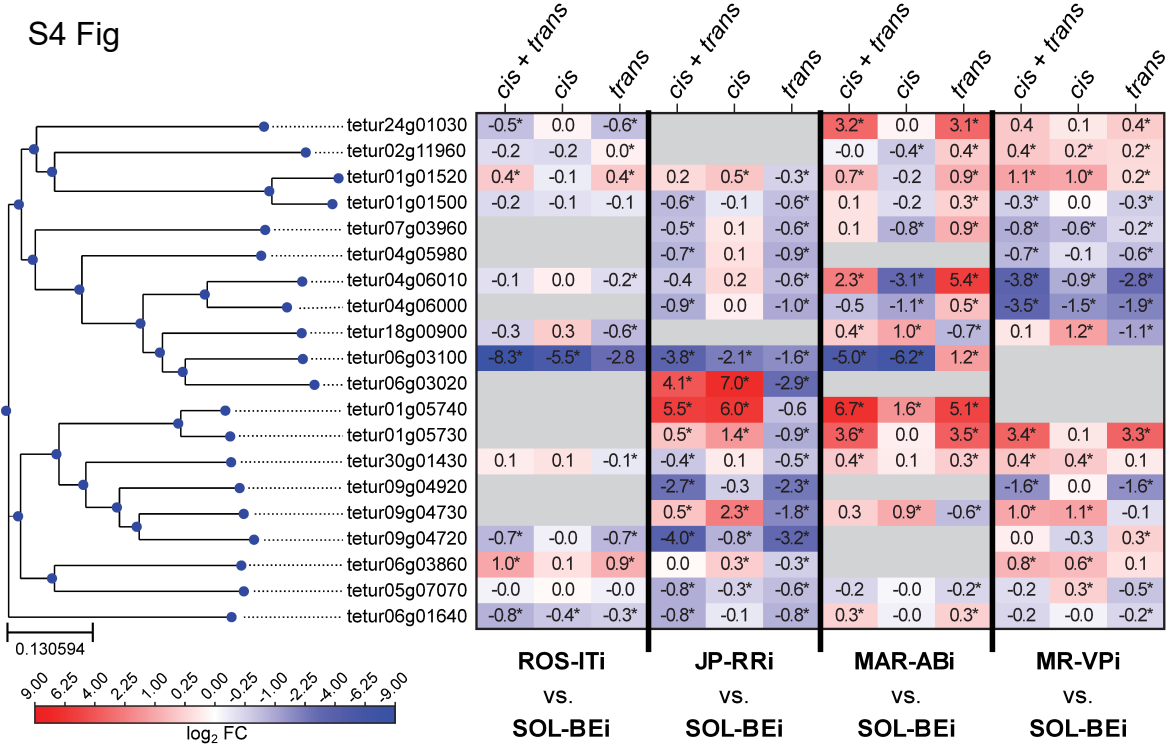
S1 Fig


A
cis
trans
synergism
antagonism
compensatory
conserved
ambiguous
All genes
D+H+P gene set
ROS-ITi
vs.
SOL-BEi
Gene number
5630
13.07%
22.04%
6.59%
4.85%
5.31%
27.71%
20.43%
44
22
26
8
19
12
62
6
44
24
61
17
ABC
CCE
CYP
dioxygenases
GST
lipocalins
MFS
PLAT
SDR
UGT
host
peptidase
B
JP-RRi
vs.
SOL-BEi
5699
10.69%
34.57%
7.77%
7.53%
5.84%
18.02%
15.58%
45
22
32
7
13
17
65
5
41
21
69
18
C
MAR-ABi
vs.
SOL-BEi
5847
11.03%
30.85%
8.76%
5.61%
6.5%
20.64%
16.61%
40
27
27
7
15
15
64
3
50
22
68
13
D
MR-VPi
vs.
SOL-BEi
5833
10.75%
31.03%
7.75%
6.22%
6.99%
21.45%
15.81%
35
27
26
9
16
17
56
6
46
25
64
18

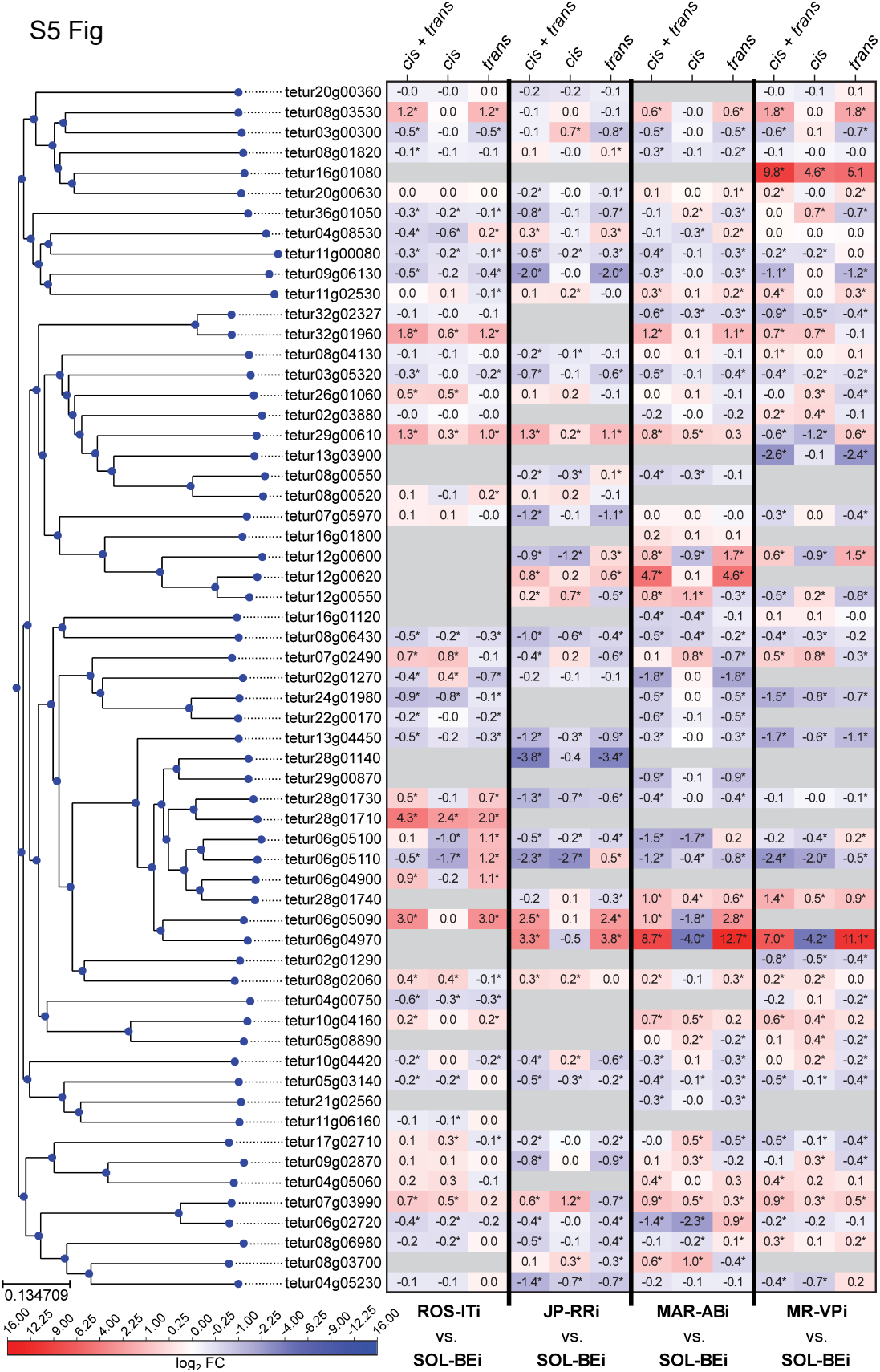
S3 Fig


cis + trans
cis
trans
DOG13
DOG8
DOG16
DOG11
DOG10
DOG1
DOG3
DOG17
DOG2
DOG15
0.130443
ROS-ITi vs. SOL-BEi
JP-RRi vs. SOL-BEi
MAR-ABi vs. SOL-BEi
MR-VPi vs. SOL-BEi
9.00 6.25 4.00 2.25 1.00 0.25 0.00 -0.25 -1.00 -2.25 -4.00 -6.25 -9.00
log2 FC

S4 Fig


cis + trans
cis
trans
tetur24g01030
tetur02g11960
tetur01g01520
tetur01g01500
tetur07g03960
tetur04g05980
tetur04g06010
tetur04g06000
tetur18g00900
tetur06g03100
tetur06g03020
tetur01g05740
tetur01g05730
tetur30g01430
tetur09g04920
tetur09g04730
tetur09g04720
tetur06g03860
tetur05g07070
tetur06g01640
0.130594
9.00 6.25 4.00 2.25 1.00 0.25 0.00 -0.25 -1.00 -2.25 -4.00 -6.25 -9.00
log2 FC
ROS-ITi vs. SOL-BEi
JP-RRi vs. SOL-BEi
MAR-ABi vs. SOL-BEi
MR-VPi vs. SOL-BEi

S5 Fig


cis + trans
cis
trans
ROS-ITi vs. SOL-BEi
JP-RRi vs. SOL-BEi
MAR-ABi vs. SOL-BEi
MR-VPi vs. SOL-BEi
tetur20g00360 -0.0 -0.0 0.0 -0.2 -0.2 -0.1 -0.0 -0.1 0.1
tetur08g03530 1.2* 0.0 1.2* -0.1 0.0 -0.1 0.6* -0.0 0.6* 1.8* 0.0 1.8*
tetur03g00300 -0.5* -0.0 -0.5* -0.1 0.7* -0.8* -0.5* -0.0 -0.5* -0.6* 0.1 -0.7*
tetur08g01820 -0.1* -0.1 -0.1 0.1 -0.0 0.1* -0.3* -0.1 -0.2* -0.1 -0.0 -0.0
tetur16g01080 9.8* 4.6* 5.1
tetur20g00630 0.0 0.0 0.0 -0.2* -0.0 -0.1* 0.1 0.0 0.1* 0.2* -0.0 0.2*
tetur36g01050 -0.3* -0.2* -0.1* -0.8* -0.1 -0.7* -0.1 0.2* -0.3* 0.0 0.7* -0.7*
tetur04g08530 -0.4* -0.6* 0.2* 0.3* -0.1 0.3* -0.1 -0.3* 0.2* 0.0 0.0 0.0
tetur11g00080 -0.3* -0.2* -0.1* -0.5* -0.2* -0.3* -0.4* -0.1 -0.3* -0.2* -0.2* 0.0
tetur09g06130 -0.5* -0.2 -0.4* -2.0* -0.0 -2.0* -0.3* -0.0 -0.3* -1.1* 0.0 -1.2*
tetur11g02530 0.0 0.1 -0.1* 0.1 0.2* -0.0 0.3* 0.1 0.2* 0.4* 0.0 0.3*
tetur32g02327 -0.1 -0.0 -0.1 -0.6* -0.3* -0.3* -0.9* -0.5* -0.4*
tetur32g01960 1.8* 0.6* 1.2* 1.2* 0.1 1.1* 0.7* 0.7* -0.1
tetur08g04130 -0.1 -0.1 -0.0 -0.2* -0.1* -0.1 0.0 0.1 -0.1 0.1* 0.0 0.1
tetur03g05320 -0.3* -0.0 -0.2* -0.7* -0.1 -0.6* -0.5* -0.1 -0.4* -0.4* -0.2* -0.2*
tetur26g01060 0.5* 0.5* -0.0 0.1 0.2 -0.1 0.0 0.1 -0.1 -0.0 0.3* -0.4*
tetur02g03880 -0.0 -0.0 -0.0 -0.2 -0.0 -0.2 0.2* 0.4* -0.1
tetur29g00610 1.3* 0.3* 1.0* 1.3* 0.2* 1.1* 0.8* 0.5* 0.3 -0.6* -1.2* 0.6*
tetur13g03900 -2.6* -0.1 -2.4*
tetur08g00550 -0.2* -0.3* 0.1* -0.4* -0.3* -0.1
tetur08g00520 0.1 -0.1 0.2* 0.1 0.2 -0.1
tetur07g05970 0.1 0.1 -0.0 -1.2* -0.1 -1.1* 0.0 0.0 -0.0 -0.3* 0.0 -0.4*
tetur16g01800 0.2 0.1 0.1
tetur12g00600 -0.9* -1.2* 0.3* 0.8* -0.9* 1.7* 0.6* -0.9* 1.5*
tetur12g00620 0.8* 0.2 0.6* 4.7* 0.1 4.6*
tetur12g00550 0.2* 0.7* -0.5* 0.8* 1.1* -0.3* -0.5* 0.2* -0.8*
tetur16g01120 -0.4* -0.4* -0.1 0.1 0.1 -0.0
tetur08g06430 -0.5* -0.2* -0.3* -1.0* -0.6* -0.4* -0.5* -0.4* -0.2* -0.4* -0.3* -0.2
tetur07g02490 0.7* 0.8* -0.1 -0.4* 0.2 -0.6* 0.1 0.8* -0.7* 0.5* 0.8* -0.3*
tetur02g01270 -0.4* 0.4* -0.7* -0.2 -0.1 -0.1 -1.8* 0.0 -1.8*
tetur24g01980 -0.9* -0.8* -0.1* -0.5* 0.0 -0.5* -1.5* -0.8* -0.7*
tetur22g00170 -0.2* -0.0 -0.2* -0.6* -0.1 -0.5*
tetur13g04450 -0.5* -0.2 -0.3* -1.2* -0.3* -0.9* -0.3* -0.0 -0.3* -1.7* -0.6* -1.1*
tetur28g01140 -3.8* -0.4 -3.4*
tetur29g00870 -0.9* -0.1 -0.9*
tetur28g01730 0.5* -0.1 0.7* -1.3* -0.7* -0.6* -0.4* -0.0 -0.4* -0.1 -0.0 -0.1*
tetur28g01710 4.3* 2.4* 2.0*
tetur06g05100 0.1 -1.0* 1.1* -0.5* -0.2* -0.4* -1.5* -1.7* 0.2 -0.2 -0.4* 0.2*
tetur06g05110 -0.5* -1.7* 1.2* -2.3* -2.7* 0.5* -1.2* -0.4* -0.8* -2.4* -2.0* -0.5*
tetur06g04900 0.9* -0.2 1.1*
tetur28g01740 -0.2 0.1 -0.3* 1.0* 0.4* 0.6* 1.4* 0.5* 0.9*
tetur06g05090 3.0* 0.0 3.0* 2.5* 0.1 2.4* 1.0* -1.8* 2.8*
tetur06g04970 3.3* -0.5 3.8* 8.7* -4.0* 12.7* 7.0* -4.2* 11.1*
tetur02g01290 -0.8* -0.5* -0.4*
tetur08g02060 0.4* 0.4* -0.1* 0.3* 0.2* 0.0 0.2* -0.1 0.3* 0.2* 0.2* 0.0
tetur04g00750 -0.6* -0.3* -0.3* -0.2 0.1 -0.2*
tetur10g04160 0.2* 0.0 0.2* 0.7* 0.5* 0.2 0.6* 0.4* 0.2
tetur05g08890 0.0 0.2* -0.2* 0.1 0.4* -0.2*
tetur10g04420 -0.2* 0.0 -0.2* -0.4* 0.2* -0.6* -0.3* 0.1 -0.3* 0.0 0.2* -0.2*
tetur05g03140 -0.2* -0.2* 0.0 -0.5* -0.3* -0.2* -0.4* -0.1* -0.3* -0.5* -0.1* -0.4*
tetur21g02560 -0.3* -0.0 -0.3*
tetur11g06160 -0.1 -0.1* 0.0
tetur17g02710 0.1 0.3* -0.1* -0.2* -0.0 -0.2* -0.0 0.5* -0.5* -0.5* -0.1* -0.4*
tetur09g02870 0.1 0.1 0.0 -0.8* 0.0 -0.9* 0.1 0.3* -0.2 -0.1 0.3* -0.4*
tetur04g05060 0.2 0.3 -0.1 0.4* 0.0 0.3 0.4* 0.2 0.1
tetur07g03990 0.7* 0.5* 0.2 0.6* 1.2* -0.7* 0.9* 0.5* 0.3* 0.9* 0.3* 0.5*
tetur06g02720 -0.4* -0.2* -0.2 -0.4* -0.0 -0.4* -1.4* -2.3* 0.9* -0.2* -0.2 -0.1
tetur08g06980 -0.2 -0.2* 0.0 -0.5* -0.1 -0.4* -0.1 -0.2* 0.1* 0.3* 0.1 0.2*
tetur08g03700 0.1 0.3* -0.3* 0.6* 1.0* -0.4*
tetur04g05230 -0.1 -0.1 0.0 -1.4* -0.7* -0.7* -0.2 -0.1 -0.1 -0.4* -0.7* 0.2
0.134709
16.00 12.25 9.00 6.25 4.00 2.25 1.00 0.25 0.00 -0.25 -1.00 -2.25 -4.00 -6.25 -9.00 -12.25 -16.00
log2 FC