# Phenotypic and Transcriptomic Responses to Stress Differ According to Population Geography in an Invasive Species

Pierre Marin[1], Angelo Jaquet[1], Justine Picarle[1], Marie Fablet[1], Vincent Merel[1], Marie-Laure Delignette-Muller[1], Mariana Galvão Ferrarini[1,2], Patricia Gibert[1,*], and Cristina Vieira[1,*]

[1]Université de Lyon, Université Lyon 1, CNRS, VetAgro Sup, Laboratoire de Biométrie et Biologie Evolutive UMR 5558, Villeurbanne, France
[2]Université de Lyon, INSA-Lyon, INRAE, BF2I, UMR0203, Villeurbanne, France

*Corresponding authors: E-mails: patricia.gibert@univ-lyon1.fr; cristina.vieira@univ-lyon1.fr.

**Accepted:** 5 September 2021

## Abstract

Adaptation to rapid environmental changes must occur within a short-time scale. In this context, studies of invasive species may provide insights into the underlying mechanisms of rapid adaptation as these species have repeatedly encountered and adapted to novel environmental conditions. We investigated how invasive and noninvasive genotypes of *Drosophila suzukii* deal with oxidative stress at the phenotypic and molecular levels. We also studied the impact of transposable element (TE) insertions on the gene expression in response to stress. Our results show that flies from invasive areas (France and the United States) live longer in natural conditions than the ones from native Japanese areas. As expected, lifespan for all genotypes was significantly reduced following exposure to paraquat, but this reduction varied among genotypes (genotype-by-environment interaction) with invasive genotypes appearing more affected by exposure than noninvasive ones. A transcriptomic analysis of genotypes upon paraquat treatment detected many genes differentially expressed (DE). Although a small core set of genes were DE in all genotypes following paraquat exposure, much of the response of each genotype was unique. Moreover, we showed that TEs were not activated after oxidative stress and DE genes were significantly depleted of TEs. In conclusion, it is likely that transcriptomic changes are involved in the rapid adaptation to local environments. We provide new evidence that in the decade since the invasion from Asia, the sampled genotypes in Europe and the United States of *D. suzukii* diverged from the ones from the native area regarding their phenotypic and genomic response to oxidative stress.

**Key words:** invasive species, *D. suzukii*, oxidative stress, transposable elements, environmental changes, genotype-by-environment interaction.

## Significance

*Drosophila suzuki*, originally from Asia, invaded the entire globe in less than 10 years. The how and why of this successful invasion are still poorly known today. The study of the response to oxidative stress for several genotypes of different geographical origin may provide us with clues to better understand this invasion story. We have identified a series of genes that are candidates for adaptation and resistance to oxidative stress, and we also show that genes that are differentially expressed following stress are more likely to lack transposable element insertions.

## Introduction

Rapid environmental changes, particularly related to human activity, can decisively affect living organisms, who must respond to them within a short-time scale (Sage 2020). Understanding the mechanisms underlying these rapid responses is challenging and could help predict organism and species survival in the face of global environmental changes (Rey et al. 2016). The rapid adaptation of invasive species to new environments, some quite different than ancestral environments, may provide insights into such

© The Author(s) 2021. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com



Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

mechanisms (Theoharides and Dukes 2007; Moran and Alexander 2014) including hormonal regulation of suites of traits, or epigenetic gene regulation (Bradshaw 1965; Ghalambor et al. 2007; Stapley et al. 2015; Marin et al. 2020). Adapting to a changing environment involves different mechanisms. Phenotypic plasticity, that is, the ability of a genotype to express different phenotypes in different environments, is a possible explanation to the success of invasive species, particularly in the case of founder populations depleted of genetic variations (Bradshaw 1965; Ghalambor et al. 2007; Chabrerie et al. 2019; Marin et al. 2020). Also, genetic diversity can rapidly increase following environmental stress if there is an activation of transposable elements (TEs) or if the epigenetic control is disturbed (Stapley et al. 2015; Rey et al. 2016; Srikant and Drost 2020). TEs, which are repeated sequences that can move around genomes (McClintock 1950), may affect the fitness of its host organism, depending on their position and interaction with the host genes. The vast majority of new TE insertions are neutral or deleterious, and purifying selection is expected to remove them or favor their silencing (Slotkin and Martienssen 2007; Fablet and Vieira 2011; Jangam et al. 2017). However, some TE insertions may be advantageous and facilitate adaptation in different environments (Naito et al. 2009; Magwire et al. 2011; Casacuberta and González 2013; Guio et al. 2014, 2018; Schrader et al. 2014; Song et al. 2015; Stapley et al. 2015; Ullastres et al. 2016; Choudhury and Parisod 2017; Horváth et al. 2017). Such adaptive effects have been previously observed in response to both biotic (e.g., virus infection) and abiotic (e.g., oxidative stress) stress (Magwire et al. 2011; Guio et al. 2014). Moreover, stress-induced changes in the epigenetic regulation of TEs, which is often sensitive to environmental cues (Spannhoff et al. 2011; Guio et al. 2018), has already been described to rapidly generate potentially advantageous changes in nearby gene regulation and facilitate rapid adaptation to environmental stress (Slotkin and Martienssen 2007; Casacuberta and González 2013).

Here, we examined the oxidative stress response of *Drosophila suzukii* flies originated from invasive and noninvasive areas, with a focus on molecular mechanisms potentially underlying the observed phenotypic differences. *Drosophila suzukii* is an Asian species of the *melanogaster* group that invaded North America and Europe in 2008 (Fraimout et al. 2017; Lavrinienko et al. 2017; Langille et al. 2017; Asplen et al. 2015; CABI). Outside of Asia, *D. suzukii* is now found in both North and South America, and throughout most of Europe, from southern Spain easterly into Poland, Ukraine, and Russia (Lavrinienko et al. 2017; Langille et al. 2017; Asplen et al. 2015; CABI). As *D. suzukii* has spread throughout the world, it has encountered and successfully colonized many different, potentially stressful environments.

To evaluate the oxidative stress response and resistance, we used paraquat. Paraquat (*N,N′-dimethyl-4,4′-bipyridinium dichloride*) is one of the most widely used herbicide in the world (Tsai 2013, 2020). Exposure to paraquat leads to the production of reactive oxygen species (ROS) and has often been used in the lab as a proxy to study oxidative stress (Rzezniczak et al. 2011; Weber et al. 2012; Oliveira et al. 2013; Deepashree et al. 2019). Resistance to oxidative stress has been associated with extended lifespan (Zou et al. 2000; Morrow et al. 2016; Deepashree et al. 2019), a trait possibly under selection during invasion of a new area. Furthermore, paraquat has been banned since 2007 in Europe but is still used in the United States and Japan.

In this study, we compared field-sampled *D. suzukii* genotypes collected in their native area of Japan with genotypes collected in invaded areas in the United States and France. For each genotype, we measured lifespan in both the presence and absence of paraquat, where we identified an effect of genotype and a genotype-by-environment interaction (GEI) effect. We went further by examining the transcriptomic response of single genotypes from each location along with analysis of TE expression. We found substantial differences among genotypes in patterns of gene expression related to oxidative stress that may explain the observed phenotypic differences and reflect population history. This work highlights the local adaptation to environmental conditions of the genotypes within a short-time scale.

## Results and Discussion

### *Drosophila suzukii* Genotypes Vary in Lifespan and Response to Oxidative Stress

Previous studies found a positive association between stress resistance and extended lifespan or aging in *Drosophila melanogaster* (Shaposhnikov et al. 2015; Deepashree et al. 2019; Moskalev et al. 2019), in which the ROS defenses are mediated by both immune and antioxidant response pathways. A similar association may be expected in *D. suzukii*, a species which diverged from the *melanogaster* group ~8 Ma. However, until now, no extensive study had been performed using *D. suzukii* wild-type genotypes.

Here, we measured lifespan and oxidative stress resistance in a total of 27 isofemale lines from six geographical regions coming from the United States, France, and Japan (supplementary table S1, Supplementary Material online). Median lifespan of flies is presented in table 1 for each population, sex and treatment and the statistical analysis of survival is presented in figure 1 and supplementary table S2, Supplementary Material online. Contrary to what has been often described in other species for sex differences in lifespan (Austad and Fischer 2016), we did not find a main effect of sex or any significant interactions with sex in our preliminary statistical model (see Materials and Methods). Therefore, male and female data were pooled for subsequent survival analysis.

In the control condition, flies from the two Japanese populations had the shortest and similar lifespan. For flies sampled

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

**Table 1**
Median Life Span (days) by Sex, Treatment, and Population

| | Japan | | France | | United States | |
|---|---|---|---|---|---|---|
| | **Sapporo** | **Tokyo** | **Montpellier** | **Paris** | **Dayton** | **Watsonville** |
| Females control | 32.0 (26.0 – 39.8) | 32.0 (29.6 – 41.0) | 40.5 (35.3 – 46.3) | 41.0 (34.5 – 49.3) | 51.4 (39.7 – 54.6) | 34.8 (27.4 – 42.4) |
| Males control | 32.6 (24.65 – 37.85) | 33.2 (27.32 – 44.43) | 39.4 (36.69 – 44.41) | 39.9 (36.48 – 48.3) | 40.4 (39.46 – 46.96) | 33.9 (26.51 – 38.92) |
| Mean control (males and females) | **32.3 ± 0.4** | **32.6 ± 0.8** | **40.0 ± 0.8** | **40.5 ± 0.8** | **45.9 ± 7.8** | **34.4 ± 0.6** |
| Females paraquat | 6.8 (5.0 – 9.1) | 5.3 (5.0 – 7.7) | 5.5 (5.2 – 6.3) | 7.1 (5.4 – 8.2) | 9.1 (8.7 – 9.8) | 5.1 (3.8 – 6.5) |
| Males paraquat | 6.1 (4.85 – 7.57) | 7.0 (5.89 – 8.2) | 7.5 (6.66 – 8.08) | 7.6 (6.81 – 9.02) | 12.0 (10.78 – 12.91) | 4.4 (3.16 – 5.17) |
| Mean paraquat (males and females) | 6.5 ± 0.5 | 6.2 ± 1.2 | 6.5 ± 1.4 | 7.4 ± 0.4 | 10.6 ± 2.1 | 4.8 ± 0.5 |

NOTE.—Values were calculated from the estimated median (see Materials and Methods) at the population level (line effects in populations were estimated as a random effect of the linear mixed model). In brackets are the 0.25 and 0.75 quantiles.

in the United States, those from Watsonville had a median lifespan very similar to the Japanese populations and were not different from the reference Sapporo population (fig. 1, value = 1.01, corresponding to ~1% greater lifespan than the reference Sapporo population). However, flies from Dayton lived the longest (value = 1.44, a 44% relative increase). The two populations collected in France lived on average 25 – 28% longer than flies in the Sapporo population (1.25 and 1.28 for Paris and Montpellier, respectively).

As expected, and in accordance to what was reported in *D. melanogaster* (Weber et al. 2012), oxidative stress had a strong negative effect on survival, with an average decrease in lifespan of 80% when paraquat was present in the medium (multiplicative coefficient of 0.20, fig. 1). Nevertheless, the decline in lifespan following paraquat treatment was variable among populations, indicating a GEI. For example, when compared with the Sapporo reference population, the flies from Tokyo, Paris, and Dayton had the same lifespan, contrary to flies from Watsonville and Montpellier, which were significantly more sensitive to paraquat treatment, with reductions in lifespan of 14.8% and 16%, respectively (multiplicative effects of $0.20 \times 0.74$ and $0.20 \times 0.80$ in fig. 1, $P$ value $< 2.03e^{-03}$ and $< 4.37e^{-02}$, respectively). Also, we observed a weak but significant correlation between the lifespan in the control and the treatment, for all genotypes ($r = 0.28$, $P$ value $< 3.3e^{04}$; supplementary fig. S1, Supplementary Material online).

In sum, Japanese flies had the lowest lifespan in the control condition but were more resistant to oxidative stress than flies from Watsonville (United States) or Montpellier (France). This GEI suggests possible local adaptation of the different populations to paraquat, perhaps associated with differences in herbicide use in the three countries. As mentioned before, paraquat is one of the most used herbicides in the world and is widely used in Japan and the United States, but forbidden in Europe since 2007 (Tsai 2013, 2020). The presence of *D. suzukii* in Europe has been reported since 2008, and flies are therefore unlikely to have encountered paraquat in the field since their arrival (Calabria et al. 2012; Asplen et al. 2015; Fraimout et al. 2017). This could have resulted in a relaxed selective pressure for oxidative stress resistance and explain why the French Montpellier population was more sensitive than Japanese and American populations (except for Watsonville). The Paris population, on the other hand, was not significantly different than the Japanese Sapporo population; this suggests an admixture event that occurred in the North of France with flies from the United States that could explain the difference between the two French populations, as previously proposed (Fraimout et al. 2017).

## Transcriptomic Variability among Genotypes

We performed a transcriptomic analysis on genotypes in somatic tissues, one from each geographical sampling location [Montpellier (MT47): France, Watsonville (W120): United States and Sapporo (S29): Japan] to identify molecular processes underlying variation in the oxidative stress response. These three genotypes are examples of what was sampled in the three areas and represent some specific local adaptation. Because they presented highly variable lifespans, we expected that the differences in the transcriptomes would be maximized. We first identified differentially expressed (DE) genes between the three genotypes in control conditions and then proceeded to the comparison of DE genes upon paraquat treatment.

A principal component analysis of gene normalized read counts (supplementary fig. S2, Supplementary Material online) showed genotype-specific clustering, independent of the paraquat treatment. To evaluate variation in the transcriptomic response of each genotype to paraquat treatment, we computed the coefficient of variation (CV) for each DE gene between control and treated flies (supplementary fig. S3,

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

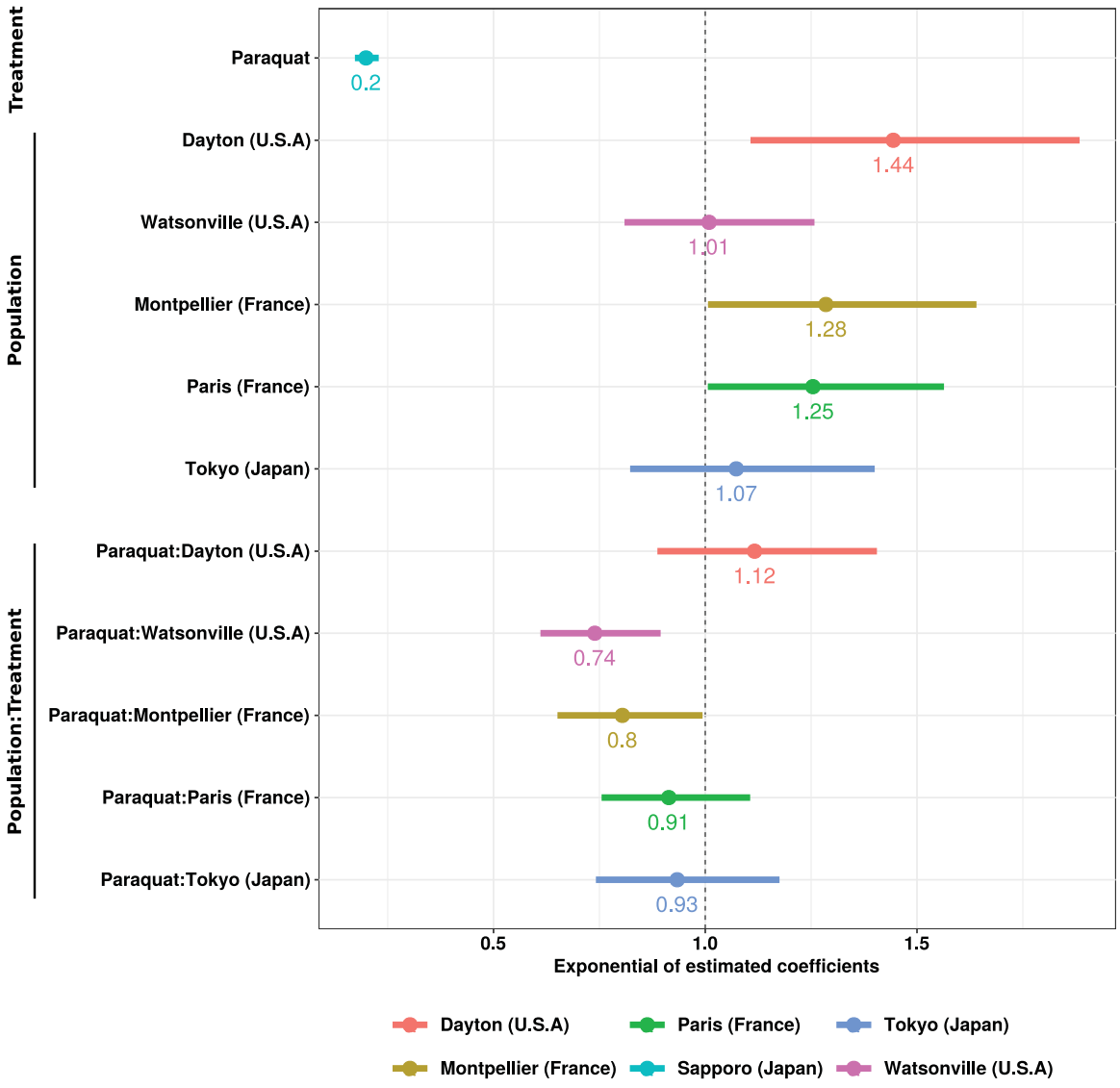


**Fig. 1.**—Lifespans under control and paraquat-treated conditions relative to the Japan Sapporo population with confidence intervals (0.95). The relative values are within each treatment group and can be interpreted as a multiplicator effect compared with the reference level (Sapporo population, female, and control condition). The intercept, which is the basal level of the Sapporo females in control condition, is equal to 31.96 days. Paraquat correspond to the mean effect of the treatment on Sapporo. Mean values of the populations correspond to the effect of the population on the lifespan compared with Sapporo in control condition and the last are the interaction term after paraquat exposure. As examples, the effect of paraquat reduced the lifespan of Sapporo to 20% (0.2) of the initial value in control condition. Values higher than 1 indicate an increase in the lifespan compared with Sapporo, whereas below 1 this indicates a higher sensitivity [e.g., Paraquat: Watsonville correspond to the interaction term and indicates that after paraquat exposure, Watsonville remains more sensitive than Sapporo with a difference of 26% (1 − 0.74)].

Supplementary Material online). CV distributions were significantly different across genotypes (paired Wilcoxon test, $P$ values $< 0.01$), which suggests significant GEI for transcriptomic response. The number of DE genes identified 1) in pairwise comparisons between genotypes in control conditions; 2) in comparisons between control and oxidative stress conditions for each genotype; and 3) in pairwise comparisons between genotypes following paraquat treatment are presented in table 2. The distribution and values of the CV were in agreement with the distribution of DE genes shown in table 2, suggesting that the difference in DE gene proportions between the genotypes were due to biological variation and not a bias of statistical power.

Due to the large number of genes that remain unannotated in the *D. suzukii* genome, a complete scenario of the genome-wide transcriptional response to oxidative stress is difficult to achieve. This may be particularly problematic when attempting to understand the functional relevance of

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

**Table 2**
Number of DE Genes between Genotypes and Treatments

| | Contrast | DE Genes | Upregulated | Downregulated | DE Proportion (%) |
|---|---|---|---|---|---|
| A | MT47\|S29 (control) | 524 | 175 | 349 | 3.6 |
| | MT47\|W120 (control) | 715 | 471 | 244 | 4.92 |
| | W120\|S29 (control) | 1023 | 208 | 815 | 7.04 |
| B | S29 (paraquat\|control) | 122 | 74 | 48 | 0.84 |
| | MT47 (paraquat\|control) | 531 | 354 | 177 | 3.65 |
| | W120 (paraquat\|control) | 281 | 214 | 67 | 1.93 |
| C | MT47\|S29 (paraquat) | 138 | 105 | 33 | 0.95 |
| | MT47\|W120 (paraquat) | 65 | 19 | 46 | 0.45 |
| | W120U\|S29 (paraquat) | 62 | 57 | 5 | 0.43 |

NOTE.—Pairwise comparisons between (A) control genotypes, (B) between treated and control flies within each genotype, and (C) in pairwise comparisons of paraquat-treated flies between different genotypes. The threshold for identifying DE genes was an adjusted *P* value $\leq 0.01$ and absolute $\log_2$-fold-change $\geq |1|$. The proportion of DE genes is the percentage of DE genes in the expressed transcriptome ($n = 14{,}538$ genes).

genotype-specific responses. However, our results suggest that the transcriptional response to oxidative stress varies across genotypes, and that some of these differences may reflect population history.

### *Basal Gene Expression Is Different between the Japanese and the Two Other Genotypes*

Pairwise comparisons of gene expression of control flies between the three genotypes revealed 715 DE genes between MT47 and W120 (4.92% of the total transcriptome), 524 between MT47 and S29 (3.6%), and 1,023 between W120 and S29 (7.04%) (table 2; supplementary fig. S4, Supplementary Material online). Most of these DE genes (~70%) had an absolute $\log_2$-fold change ($\log_2$FC) below 2 (supplementary fig. S4, Supplementary Material online) and only 60 had an absolute $\log_2$FC higher than 5. In agreement with our results, in *D. melanogaster* the genotypic differences also accounted for ~7% of DE genes showing microenvironment plasticity among 16 DGRP (Drosophila Genetic Reference Panel) lines reared under carefully controlled standard conditions (Lin et al. 2016).

Most of the DE genes correspond to biological processes such as metabolism or protein synthesis and may possibly reflect genotype-specific differences related to local adaptation. In general, the level of expression for DE genes in the United States (W120) and France (MT47) genotypes was lower than in the genotype from Japan (S29), suggesting the S29 genotype has by default a higher level of transcription for the DE genes.

To further examine these DE genes, we performed a Gene Ontology (GO) analysis with genes that had an ortholog in *D. melanogaster* (supplementary fig. S5, Supplementary Material online). The rationale was to identify transcriptomic differences possibly related to adaptation of the different genotypes to their respective environments. We detected 44 GO terms shared between MT47 and W120 in comparison with S29. These terms were mainly related to translation, protein metabolic process, ribosome biogenesis, response to hyperoxia, and immune response (antibacterial related). These terms were downregulated in W120 or MT47 genotypes when compared with S29. We also detected other functional terms in molecular function that seemed to be specifically downregulated in W120 (so they appear in both W120 vs S29 and W120 vs MT47 results): carbohydrate transport and energy metabolism. It is plausible to say that these functions are modified in W120.

### *Oxidative Stress Induces Gene Upregulation in MT47 and W120 Genotypes*

We also compared changes in gene expression between flies in control conditions versus 10 mM concentration of paraquat and identified a total of 659 unique DE genes across the 3 genotypes (fig. 2 and table 2). Interestingly, the proportion of the transcriptome affected by oxidative stress differed between the Japanese (S29) and the two other genotypes (MT47 and W120), and the Japanese genotype appeared highly stable, with fewer DE genes in response to paraquat (122 genes, representing 1.10% of the transcriptome). We detected 281 DE genes in W120 (2.46%) and 531 in MT47 (4.51%). In this way, exposure to paraquat affected the expression of up to 5% of the transcriptome of *D. suzukii* (659 DE genes between control and paraquat) with a majority of DE genes being upregulated (435/659). Similar changes in gene expression have been observed in *D. melanogaster*, with 608−1,111 DE genes identified after exposure to 5 or 15 mM of paraquat, respectively (Girardot et al. 2004). The comparison between MT47 and W120 showed that a greater number of DE genes were uniquely shared (114) between these two genotypes. It should be pointed out that these two genotypes were the most affected by paraquat exposure in our phenotypic analysis (fig. 1), suggesting that stress sensitivity could be linked to greater transcriptional deregulation. The French genotype had by far the greatest number of DE genes, consistent with this hypothesis of transcriptional

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

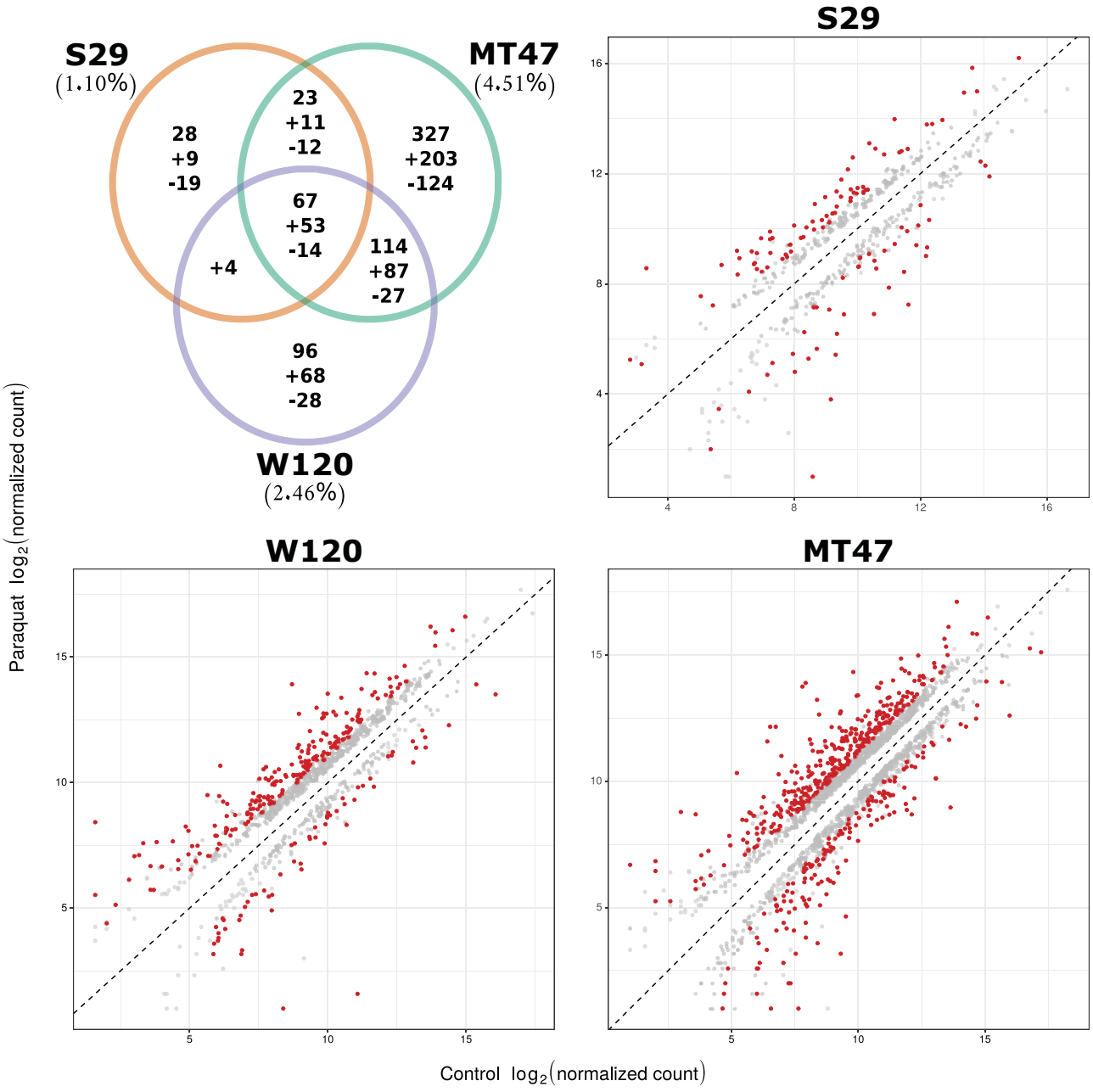


**Fig. 2.**—Gene expression between control and paraquat conditions. A Venn diagram of shared and unique DE genes identified in comparisons of paraquat-treated and control flies within each population after paraquat exposure (top left). Scatter plots of $\log_2$ normalized read count for Japan (top right), United States, and France (bottom panels) comparing control and treated flies. Individual genes are indicated by dots. Red color corresponds to significant DE genes (see Materials and Methods).

deregulation, and could reflect a lack of adaptation to paraquat, which has been banned as an herbicide in Europe since 2007, prior to the arrival of *D. suzukii*.

Out of the 659 DE genes, we also performed a GO enrichment analysis for each genotype (paraquat vs control) with 621 orthologous genes in *D. melanogaster*. We were able to detect enriched terms in downregulated genes from S29 and in upregulated genes from W120 and MT47. These observations are in accordance with the fact that a functional major upregulation of genes in response to paraquat was observed only in W120 and MT47. When comparing the GO terms enriched in upregulated genes (supplementary fig. S6, Supplementary Material online), terms such as ligase activity, oxidation–reduction, ATP binding, drug binding, and ion binding were common to W120 and MT47. As observed in related species,

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

paraquat can indeed cause DNA damage via oxidative stress (Rzezniczak et al. 2011). MT47 had a greater number of specific enriched terms, mostly related to DNA repair (including aforementioned ligase activity and telomere maintenance, among others), protein translation, protein refolding, and mitochondrion. W120 had other enriched terms related to carbohydrate metabolism, and a surprising result was that we identified an enrichment in terms related to copper detoxification and response to metal ions that were not found in the other genotypes. Previous studies made on bacteria and tissue culture have demonstrated a tradeoff between copper tolerance and sensitivity to paraquat (Kohen et al. 1985; Zer et al. 1991). There were no enriched GO terms among upregulated genes in S29 and downregulated genes were mainly related to immune response (to bacteria), response to increased oxygen levels (hyperoxia) and peptidase activity (supplementary fig. S6, Supplementary Material online). Overall, it appeared that although paraquat induced increased expression for genes related to oxidation−reduction, detoxification, drug/metal binding, DNA repair, and protein refolding in MT47 and W120, it reduced the expression of important genes for the antioxidant response in S29.

This dissimilarity between the Japanese and the other two genotypes was seen in three levels: First, the S29 genotype had the lowest lifespan under standardized control conditions but was one of the genotypes most resistant to oxidative stress; second, under control conditions, this genotype had the highest amount of DE genes when compared with MT47 and W120; and third, the transcriptional response to paraquat involved a much smaller portion of the genome and there were fewer DE genes unique to this genotype. Together, these results suggest that the Japanese genotype maintained some constitutive defenses to oxidative stress. In the absence of oxidative stress, the expression of constitutive defense may come at the cost of reduced lifespan but it would result in greater resistance when flies encounter paraquat. In the case of the French and American genotypes, many upregulated genes are directly related to the oxidative stress response (GO enriched for oxidation−reduction, immune response and ion binding), which could indicate they are experiencing a greater amount of oxidative damage and therefore explain why they were more affected by paraquat exposure.

A previous study using a large sample of *D. suzukii* populations all over the world to perform a genome scan identified 26 gene candidates to be involved in the invasive process (Olazcuaga et al. 2020). None of them was found in our data set of DE genes. This can be due to the stringent thresholds we used to consider a gene as DE, but also to the fact that we are using only three genotypes, which may be not enough to identify these genes.

### *DE Genes Common to the Three Genotypes Were Mostly Upregulated with Oxidative Stress*

From the 659 DE genes between control and paraquat exposure, 67 were shared by all genotypes. This set of core genes were regulated in the same way for the three genotypes: 14 downregulated (from $\log_2FC = -1.03$ to $-10.8$) and 53 upregulated (from $\log_2FC = 1.03-10.48$) (fig. 3).

Among those upregulated following paraquat treatment, we found genes related to stress response such as *Hsp* and *Cyp* genes families. Among the strongest upregulated genes we detected a gene encoding for a transcription factor (TF) A ($\log_2FC = 10$) and other genes in the *Hsp* gene family. This set of common DE genes likely corresponded to those directly implicated in stress response. Among the strongest downregulated genes, we identified a cytochrome P450 gene that was the top downregulated gene in all three genotypes and has a $\log_2FC < -10$ in the Japanese genotype.

We performed a GO enrichment analysis for the set of 67 genes common to the three genotypes. Out of 53 upregulated genes, 37 had homologs in *D. melanogaster*. Enrichment analysis on this set of genes identified only one significant GO term: ligase activity (which is related to DNA repair). Four of the five genes within this GO term were tRNA-ligases, which may play a role in protecting cells against oxidative damage following their translocation into the nucleus (Czech et al. 2013). For downregulated genes, because only 9 of the 14 genes had a homolog in *D. melanogaster*, we manually verified GO terms related to each gene. Interesting terms associated with peptidoglycan metabolic process and negative regulation of NK cell differentiation involved in the immune response were related to two specific genes: *PGRP-SC1a* and *PGRP-SC1b*. *PGRPs* (Peptidoglycan recognition proteins) are important in recognizing and degrading bacterial peptidoglycan, although *PGRP-SC1b* has not shown antibacterial activity and may instead be a scavenger protein (Bischoff et al. 2006).

### *Regulatory Network Analysis on Genes Presenting GEI Reveals a Module Involved in Stress Response and Adult Lifespan*

Genes presenting GEI represent the set of DE genes with expression differentially modulated by oxidative stress according to genotype (fig. 4 and table 2C), and may provide insights into the mechanisms of local adaptation (Lasky et al. 2018). We identified 52 genes in which the GEI was driven by a differential response in only one genotype compared with the other two. This included 22 genes differentially modulated in MT47 compared with the others, 14 in S29 compared with MT47 and W120, and 16 for W120 against others. We have presented some examples that highlight the potential effect of genotype-specific responses to oxidative stress (fig. 4; supplementary table S3, Supplementary Material online). These genes were chosen for the greatest $\log_2FC$ and

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

**Fig. 3.**—DE genes (67) shared by all genotypes after paraquat exposure (W120 in red, MT47 in green and S29 in blue). The plotted labels refer to a TE insertion detected within a gene (intron, exon, 5′, or 3′ UTR) or within the 2 kb flanking regions of a gene.

illustrate cases in which the magnitude of the response to paraquat differed among genotypes. Interestingly, the French genotype showed massive upregulation of some *Hsp* genes compared with the American and Japanese genotypes. These genes are known to be highly responsive to temperature (Chen et al. 2015) and also to oxidative stress (see review, King and MacRae 2015).

One caveat of genome-wide expression analysis is the statistical power to identify biologically relevant differences in expression levels. We have applied a threshold (FDR $< 0.01$ and absolute $\log_2$FC $> 1$) to identify DE genes, but it is possible that genes with more subtle changes in expression are also important. Indeed, genes showing evidence of GEI are often found in upstream parts of regulatory networks, where

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

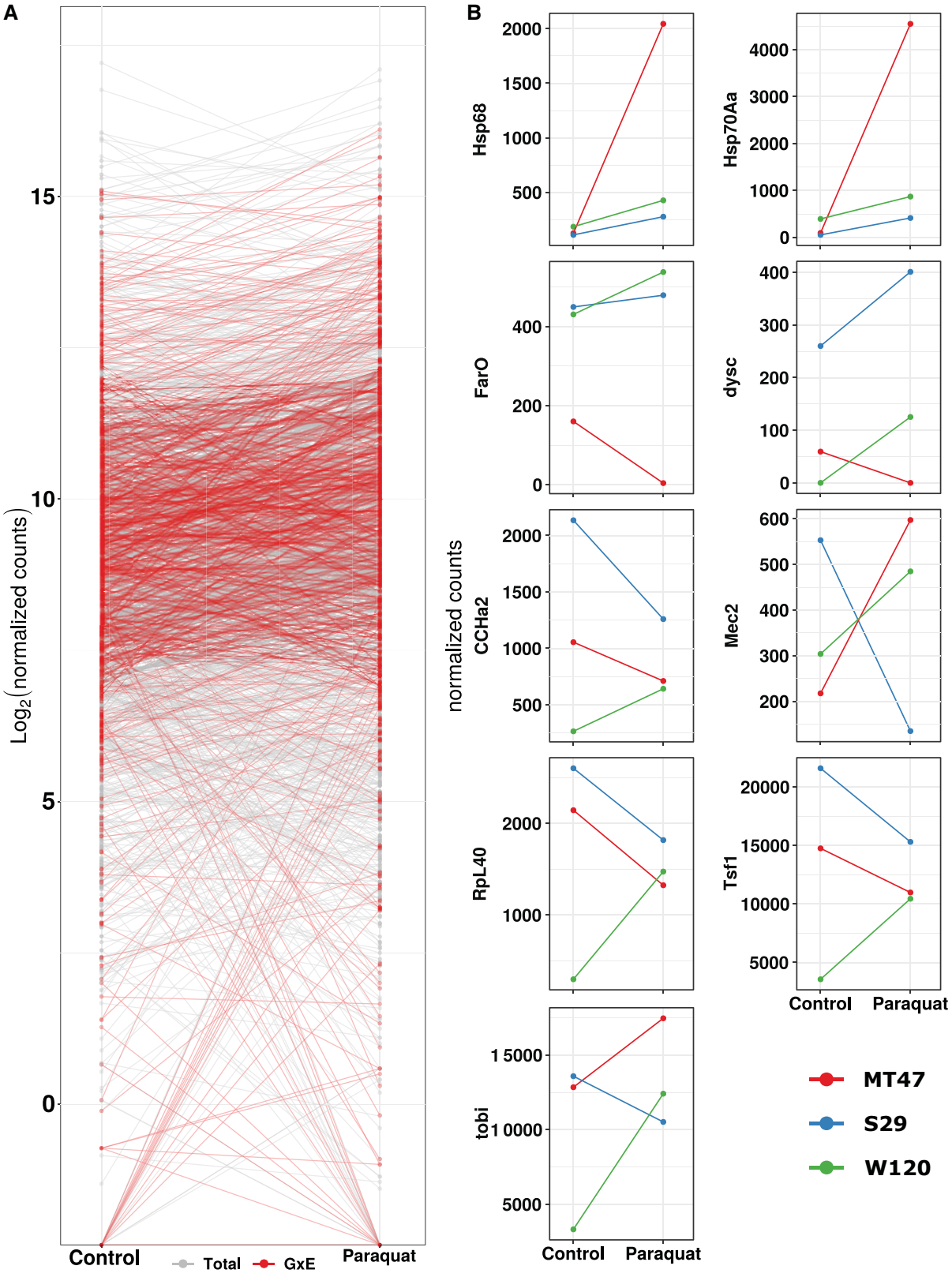


FIG. 4.—Reaction norms for DE genes between control and paraquat using log2 of normalized counts. (*A*) Reaction norms of all DE genes, in red for GEI genes, and in gray for DE genes without GEI. (*B*) Examples of 9 DE genes with GEI, with colors referring to genotype (red for MT47, blue for S29, and green for W120).

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

even very small differences in expression could have pronounced phenotypic consequences.

In this way, in order to gain more information on the 213 genes presenting a GEI and their regulation, we predicted a regulatory network between these and putative TFs. Based on the complete set of TPM normalized counts, we calculated using the GReNaDIne method (Peignier et al. 2020) the *Z*-scores for each gene to have a relative expression per gene and per sample. Putative TFs (supplementary table S4, Supplementary Material online) were recovered based on homolog genes to annotated TFs from *D. melanogaster*. The *Z*-scores were used to predict gene regulatory networks between TFs and genes by using two methods: Genie3 (Huynh-Thu et al. 2010) and Pearson's correlation. We selected only genes and TF pairs predicted by both methods. Finally, from the network visualization (supplementary fig. S7*A*, Supplementary Material online), we detected a highly connected module that accounted for 123 GEI genes and 17 TFs (supplementary fig. S7*B*, Supplementary Material online), mostly expressed in paraquat-treated flies from MT47 and W120 in comparison to S29. The most connected TF detected was *SREBP* (43 edges), followed by *stc* (27 edges), *Xbp1* (26 edges), *Ets21C* (25 edges), *Sox21a* (24 edges), and *Dr* (19 edges); other TFs within this module were *sima*, *nub*, *CG1529*, *cnc*, *MTF-1*, *HLH54F*, *Atf3*, *vnd*, *CHES-1-like*, *CG5694*, and *Alh*. GO analysis of this highly connected module recovered an enrichment related to protein refolding, response to stress and hypoxia, and more interestingly a 5.67-fold enrichment (FDR-adjusted *P* value = 0.0481) of the term *determination of adult lifespan* (GO: 0008340).

## TE Dynamics upon Oxidative Stress

TEs have been described as stress sensitive and their activation by stress-responsive elements in promoting regions can generate a burst of transposition and facilitate adaptation by increasing the genetic diversity upon which selection acts (Casacuberta and González 2013; Dubin et al. 2018). A recent review cited several examples of TE family activation following stress, which may depend on the type of stress and the TE family (Guerreiro 2012; Horváth et al. 2017). Moreover, TE transcription is a prerequisite to TE activity and environmental changes can affect the expression of TEs by lifting epigenetic repressive regulation mechanisms (Fablet and Vieira 2011). In this context, we studied the dynamics of TE expression as well as TE insertion near DE genes to detect TE-related regulatory mechanisms.

### *TE Expression Is Not Sensitive to Oxidative Stress in D. suzukii*

In our experiments, TE expression levels were very low, and reads corresponding to TEs did not exceed 3.8−7.1% of the total transcriptome. In control condition, DETEs identified in pairwise comparisons between genotypes represented from 3.08 (48 families) to 5.91% (92) of the total number of TE families annotated in the *D. suzukii* genome (supplementary table S5, Supplementary Material online). W120 exhibited a greater level of DETEs compared with other genotypes, with almost 70 TE families upregulated. Moreover, a similar number of both up- and downregulated DETEs were identified in the comparison between MT47 and S29 (fig. 5).

After paraquat exposure, very few TE families changed in their expression levels (supplementary table S5, Supplementary Material online). In total, only 12 TE families were DE (fig. 5). Six TE families in MT47 and three in S29 were upregulated. In W120, differential expression of five TE families was observed, with three showing upregulation and the remainder, downregulation. Among the DETEs, all classes of TE families were represented. We observed a differential expression in two of the genotypes MT47 and W120 in a *Copia* cluster and a *Tc1* mariner cluster, which could suggest specific activation of these TE families upon oxidative stress.

In conclusion, we show that in *D. suzukii* very few TEs were activated after oxidative stress, with a maximum of six TE families deregulated following exposure to paraquat in the French genotype. This result is not related to the potential activity of TEs in *D. suzukii*, because a greater number of TE families are DE between genotypes in control conditions, suggesting that TEs in *D. suzukii* are capable of being not only expressed, but also DE.

### *TE Insertions Near DE and GEI Genes*

Most TE insertions are neutral or slightly deleterious, but some may be beneficial and implicated in adaptation (Casacuberta and González 2013; Rey et al. 2016; Dubin et al. 2018; Mérel et al. 2021). The impact of TE insertions is often achieved by their effect on gene expression, likely due to the addition of regulatory sequences, present within TEs, particularly during stress (Rebollo et al. 2012). Although TEs have been suggested as playing an important role in the success of invasive species, by generating genetic diversity and thus compensating for bottleneck effects after introduction, no empirical data exists to support this hypothesis (Schrader et al. 2014; Marin et al. 2020; Mérel et al. 2021).

As TEs represent ~47% of the *D. suzukii* genome (Mérel et al. 2021) and can potentially interfere with gene expression during stress (Guio et al. 2018), we tested the hypothesis that they could modulate gene expression of nearby genes by the addition of regulatory regions.

To test for an enrichment or depletion of TE insertions inside and around DE genes, we first tested if the distribution of TEs in the three genomes was not significantly different. Our results show that the distribution of insertions along the chromosomes did not differ among the three genomes (chi-square test = 0.69; supplementary table S6, Supplementary Material online), and, as observed in other *Drosophila* species, a

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

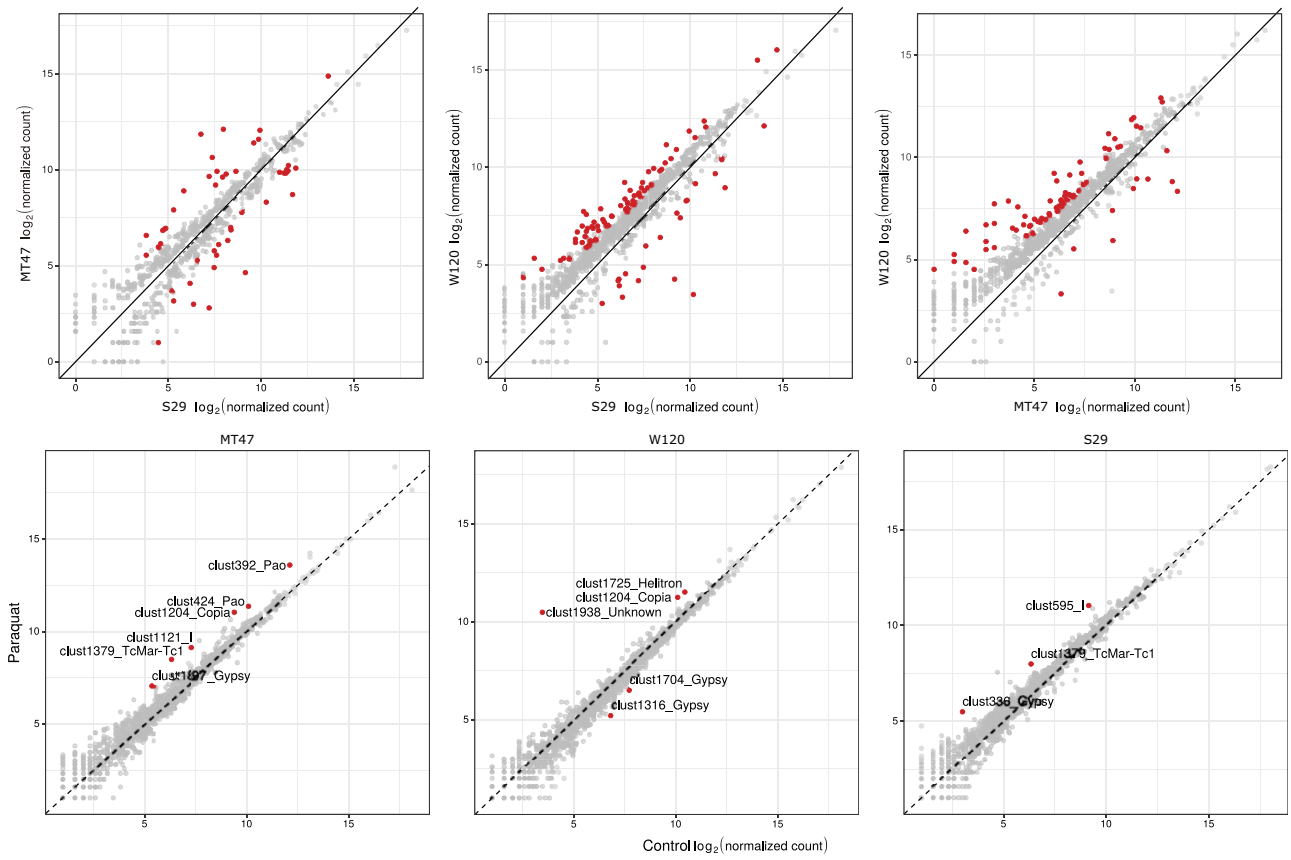


**Fig. 5.**—TE expression between genotypes (upper panel) and between control and paraquat conditions for each genotype (bottom panel). Scatter plots represent the $\log_2$ normalized read counts. Individual TEs are indicated by dots. Red color corresponds to significant DE TEs (see Materials and Methods).

majority of the insertions were in intergenic and intronic regions (Kaminker et al. 2002).

We then tested the dependence of TE insertions and gene expression states (DE or not) after paraquat exposure (supplementary table S7, Supplementary Material online). Chi-square tests for the three genotypes showed that DE genes had fewer than expected TE insertions within genes or the 2 kb flanking region (*P* values are $2.16^{e-03}$ for MT47, $1.32^{e-02}$ for S29, and $2.6^{e-02}$ for W120).

This observed depletion of TE insertions near DE genes suggests that TE insertions in stress response genes may be eliminated by strong purifying selection. This paucity of TE insertions was also observed for DE genes that were shared by the three genotypes: Of the 67 shared DE genes that responded to paraquat treatment in similar ways across the three genotypes (fig. 3), we found 11 genes with one or more TE insertions. A gene encoding a glutathione-*s*-transferase was the only one to display a common TE insertion in its flanking region. Finally, we tested for the enrichment of TE insertions inside or in the near 2 kb in genes presenting GEI. We found a total of 53 genes with at least 1 TE insertion. In general, the DE genes presenting a GEI, had the same distribution of TE insertion as all genes, except for genes with a GEI between MT47and W120 ($P$ value $< 0.016$; supplementary table S7, Supplementary Material online). Also, the TE insertions were associated with up- or downregulation, with no general pattern (fig. 6; supplementary tables S8 and S9, Supplementary Material online summarize detected TE insertions in GEI genes for the different genotypes). In a previous study by Mérel et al. (2021), the authors identified several TE insertions nearby genes that could have been associated with the invasive process. From the 15 genes with TE insertions detected by the authors in that study, none of them were identified in our study. One possible reason for this is that we are specifically looking for genes that respond to oxidative stress, which is not the case in the previous study.

Moreover, several insertions were found in 3′ and 5′ UTR and could have regulatory impacts on those genes due to TF-binding motifs (TFBMs) TEs harbor. We investigated whether this was the case in our data set and focused on three TFBMs from proteins that belong to the antioxidant response elements (ARE) group (Villanueva-Cañas et al. 2019): CnC (Cap'n'collar), Hsf (Heat shock factor), and dl (dorsal).

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

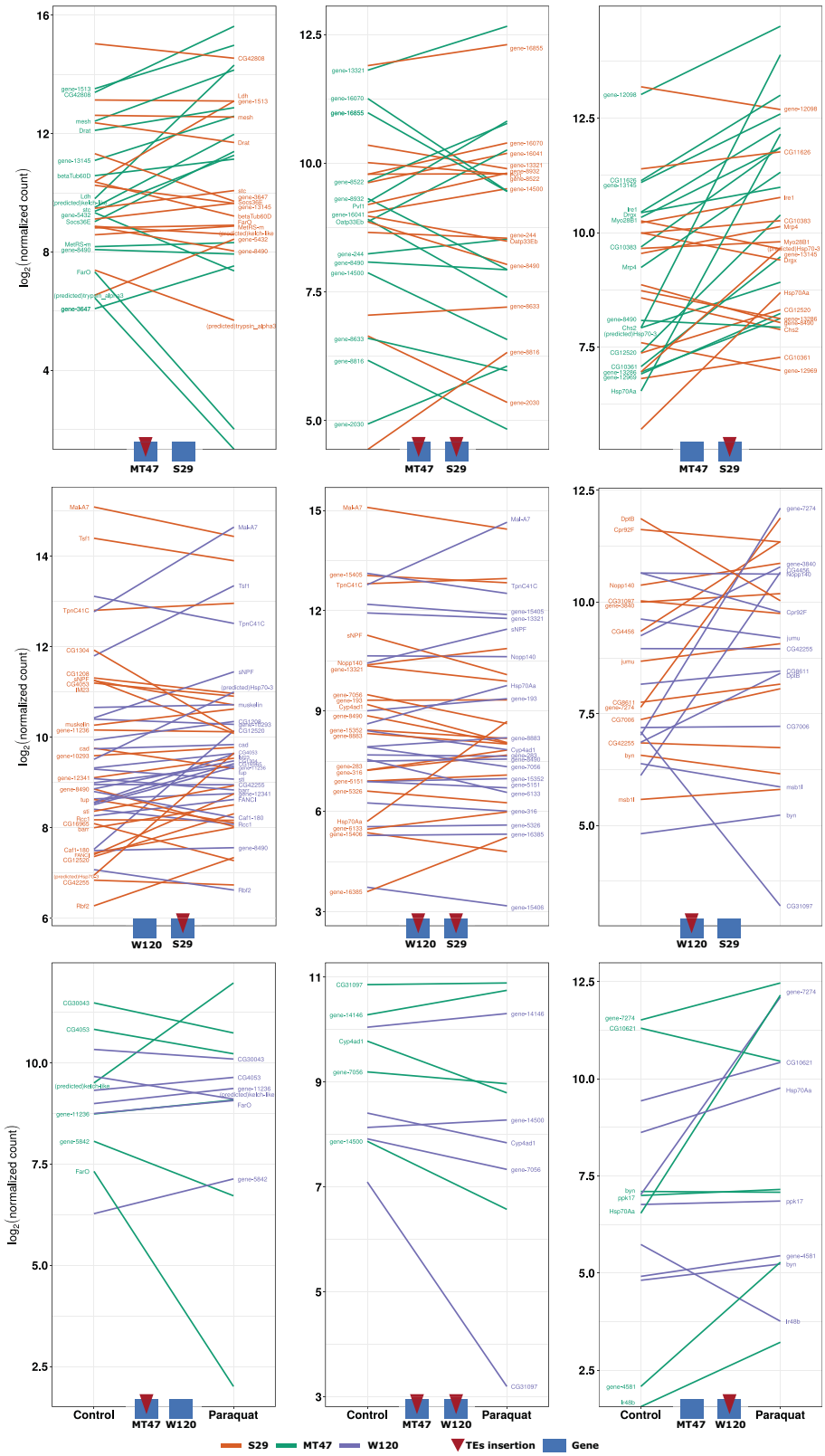


**FIG. 6.**—Expression reaction norms (in $\log_2$-fold-change) for DE expressed genes showing evidence of GEI. In each panel are represented DE genes for pairs of genotypes presenting GEI. For each comparison, in the bottom of the panel, is represented the status of the genes regarding the presence or absence of a TE insertion (red triangle). Each genotype is represented by a color.

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

Interestingly, the *cnc* gene did not present a strong GEI (padj $< 0.05$; $\log_2$fc $\sim 0.38$ between MT47 and S29), but was detected as a central node within our GEI regulatory network. This is in accordance with the idea that most TFs are low abundance essential genes from prokaryotes to higher eukaryotes, but even small changes in TFs might impact the expression of regulated genes (Janga and Contreras-Moreira 2010; Lambert et al. 2018; Pokhilko et al. 2021). In this way, we analyzed enrichment for the three aforementioned TFBMs and all TE families identified near DE genes (supplementary fig. S8*A*, Supplementary Material online). Of 196 TE families, 31 had at least one TFBM, most of them related to CnC and Hsf. We also found seven GEI genes with TFBM linked to a TE insertion (supplementary fig. S8*B*, Supplementary Material online), including the *stc* gene (*shuttle craft*), encoding an NFX1 family TF implicated in modulating adult lifespan and aging (Pasyukova et al. 2004), and which was also detected within our regulatory network related to oxidative stress and lifespan.

Even though further analyses are needed to understand the molecular mechanisms responsible for changes in gene expression for this category of genes, we believe we have provided new evidence on the putative regulatory effects of TEs in response to oxidative stress in *D. suzukii*.

## Conclusion

Our results showed a difference in oxidative stress response between native and invasive populations of *D. suzukii*, that was not homogeneous between sampling sites in the same country. The differences observed between the two French populations could be explained by differential admixture subsequent to colonization in these two regions of France. In the United States, possible local adaptation to copper in the environment in Watsonville, as revealed by the functional analysis, may explain the observed difference. Further research is required to test these hypotheses. Our data also reveal that gene expression patterns first depend on the genotype, and on the stress condition to a lesser extent. By regulatory network analysis, we also detected interesting TFs that might be playing differential roles in each genotype. Finally, we showed that contrary to expectations, oxidative stress does not induce significant activation of TEs and that DE genes under stress conditions are depleted of TE insertions in the three genotypes of *D. suzukii* studied. Our results highlight that it is important to focus on several genotypes in performing phenotypic or transcriptomic analysis, and that we should consider the neglected role of TEs in adaptive evolution. Also, we provide a good example of how phenotypic and molecular approaches should complement each other to better understand the evolution of biological traits.

## Materials and Methods

### *Drosophila suzukii* Lines Rearing Conditions and Phenotyping

*Drosophila suzukii* genotypes were sampled in 2014 in the native area (Japan: Sapporo and Tokyo) and two invaded areas (United States: Watsonville and Dayton and France: Montpellier and Paris) (supplementary table S1, Supplementary Material online). To establish isofemale lines, a single gravid female was placed in a culture vial, and the line maintained thereafter with a low larval density in vials containing modified "Dalton" medium (supplementary table S10, Supplementary Material online) in a controlled environment: $22.5 \pm 1$ °C, 70% $\pm$ 5% RH (relative hygrometry) and a 16:8 (Light/Day) (Dalton et al. 2011).

We used paraquat (methyl viologen dichloride hydrate, ref. 75365-73-0, Sigma-Aldrich) to mimic oxidative stress. Paraquat (10 mM) was added to the cooling medium, before pouring into vials. Control vials were made at the same time but without adding paraquat. In the experiment, ten 4- to 7-day old flies were placed in experimental vials and transferred to new vials every $3-4$ days to limit microbial development. Both males and females were tested and kept in separate vials. Survival was monitored by visual inspection every 24 h. There were three replicate vials for each combination of the 27 isofemale lines, sex, and paraquat treatment, for a total of 324 vials (supplementary tables S1 and S11, Supplementary Material online).

### Survival Data Analysis

The analysis of survival data was performed in two steps on *R* software (*v.3.6.0*; R Core Team 2019). First, for each replicate (ten survival times), we used the fitdistcens function from the *fitdistrplus* package (*v.1.0-14*), to determine which of several distribution models (Weibull, lognormal, and gamma) were most appropriate to fit our right censored (33 flies) data (Delignette-Muller and Dutang 2015). The Weibull distribution was chosen after graphical comparison with others, also confirmed using loglikelihoods of the fitted models. For each replicate, the fitted distribution was summarized using its theoretical median. Second, a linear mixed model was fitted to the log-transformed medians using the lmer function of the *lme4 R* package (*v.1.1-21*; Bates et al. 2015), and *P* values were estimated using *lmerTest* (*v.3.1-0*; Kuznetsova et al. 2017) with treatment, sex and population (the six sampled cities) entered as fixed factors and isofemale line as a random factor. The main effect of sex and interactions with both treatment and population were removed after AIC comparison from the final model for analysis. The interaction between population and treatment (GEI effect) was kept in the model. Model coefficients are reported with their confidence intervals (0.95) in supplementary table S2, Supplementary Material online, and after exponential transformation in figure 1. These

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

effects can be interpreted as a multiplicative effect on the median lifespan compared with a reference, here chosen as the nonexposed group from Sapporo. So, for example, with the control Sapporo flies centered on 1, an effect of 0.2 for paraquat-treated Sapporo flies means they have 20% of the survival time of Sapporo flies without paraquat. Normality and homoscedasticity of residuals and normality of random effects were confirmed graphically after logarithmic transformation of median survival times. We also examined the correlation across the isofemale lines between log-transformed survival times in control and paraquat-treated conditions using a Pearson correlation coefficient in R (supplementary fig. S1, Supplementary Material online).

## DNA Extraction and Sequencing

We sequenced genomic DNA for one isofemale line per country: S29, W120, and MT47, respectively, from Sapporo (Japan), Watsonville (United States) and Montpellier (France). DNA was extracted using phenol chloroform from a pool of ten adult females. Libraries and sequencing were performed by the platform GeT-PlaGe, Génopole Toulouse/Midi-Pyrénées (France), using Illumina (150 bp) *TruSeq Nano pair end*. We obtained between 33,362,864 and 72,022,388 reads per library. Sequences were cleaned using *Trimmomatic* with default parameters (Bolger et al. 2014).

## RNA Extraction and Sequencing

We used the same three isofemale lines (S29, W120, and MT47) for our analysis of gene expression. For each of two biological replicates, fifteen 4- to 7-day-old-females were exposed for 24 h to medium supplemented with paraquat (20 mM) or without paraquat (i.e., a total of 12 samples). Flies were dissected on ice in a phosphate buffer saline solution to remove gonads, and the remaining somatic tissue was frozen in liquid nitrogen and stored at −80 °C.

We used the RNAeasy Plus Mini Kit (Qiagen) to extract total RNA from the somatic tissues, following the protocol provided by manufacturer. Samples were treated with DNAse (ref. AM2224, AMbion) according to the manufacturer's instructions and stored at −80 °C. RNA amount and quality was estimated using Qubit (Thermo Fisher Scientific) and the 2100 Bioanalyser instrument (Agilent). RNA libraries and sequencing were performed on the GenomEast platform, a member of the 'France Génomique' consortium (ANR-10-INBS-0009). Libraries were constructed using the TruSeq Stranded mRNA Library Prep Kit following the manufacturer's recommendations. The libraries were sequenced on Illumina High HiSeq 4000 with paired-end 100 base pair long reads.

## Transcriptome Analysis

Between 62.76 and 120.12 million paired-end reads were generated from the 12 libraries. Quality was assessed using *FastQC* (*v. 0.10.1*), a trimming step implemented with *UrQt* (*v. 1.0.17*, minimum phred score of 20), and quality was checked again using *FastQC* (Modolo and Lerat 2015). RNA-seq data were mapped on the *D. suzukii* reference genome using *HISAT2* (*v. 2-2.1.0*) and read counts for genes (supplementary table S12, Supplementary Material online) were computed with *eXpress* (Roberts and Pachter 2013; Kim et al. 2019; Paris et al. 2020). We performed a reciprocal *BLASTN* (v. *2.2.26*) between the *D. suzukii* genes and the *D. melanogaster* database (FlyBase, dm6 version) (archive data: FB2018_06) in order to identify orthologs (Altschul et al. 1990). Another *BLASTX* was performed against the NCBI nr database, using predicted genes *in D. suzukii* for which no orthologues were detected in *D. melanogaster*. Matched hits from this *BLASTX* were tagged with the term "(predicted)." Of the 16,905 annotated genes in the *D. suzukii* genome, 8,428 matched with a Flybase gene and 478 others on the nr database (52.7% of total genes).

Differential expression analysis was performed using the *DESeq2* package (*v. 1.24.0*) on *R* (*v. 3.6.0*) (Love et al. 2014). We built a model estimating the effects of genotype (MT47, W120, and S29), the environment (control and paraquat), and the GEI effect. The *lfcShrink* function was used to estimate $\log_2$-fold change and identify DE genes using the *ashr R* package (Stephens 2017). DE genes (supplementary table S13, Supplementary Material online) were those with an FDR-adjusted *P* value below 0.01 and absolute $\log_2$-fold change $>1$. The CV (standard deviation/mean) on normalized counts was computed for each genotype, between control and paraquat.

## Functional Enrichment Analyses

We performed a GO enrichment analysis (supplementary tables S14−S16, Supplementary Material online) directly on the geneontology.org website, using homologs in *D. melanogaster* to discover over or under-represented gene functions from the lists of DE genes (Gibson 2008). *P* values were calculated using a Fisher test for enriched GO terms and adjusted with the Benjamini−Hochberg correction for multitesting (Benjamini and Hochberg 1995). GO terms with FDR $\leq 0.05$ were defined as significantly enriched. The GO terms were reduced to representative nonredundant terms using the *REVIGO* tool and manual curation (Supek et al. 2011).

## Regulatory Network Prediction

We predicted a regulatory network with the *GReNaDIne* method (Peignier et al. 2020) between putative TFs and genes presenting a GEI. From the complete set of TPM normalized counts for all transcripts, we calculated *Z*-scores for each gene in order to have a relative expression per gene and per sample. Normalized *Z*-scores represent the number of standard-deviations that a given value is away from the mean of all

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

the values in the same group (each gene for all conditions). Putative TFs (supplementary table S4, Supplementary Material online) were recovered based on homologue genes to *D. melanogaster* information available at the gene ontology level (term "DNA-binding TF activity," GO:0003700) along with information from *Jaspar* (Fornes et al. 2020). The *Z*-scores were used to predict gene regulatory networks between TFs and GEI genes by using two methods: Genie3 (Huynh-Thu et al. 2010) and Pearson's correlation. We filtered the 10k best interactions for each method and selected only genes and TF pairs that were predicted by both methods. Finally, we generated a network visualization with *cytoscape* (v3.6.1) (Shannon et al. 2003) and were able to detect highly connected components and highly connected nodes for further investigation by GO enrichment analysis as explained above (supplementary table S17, Supplementary Material online).

### TE Identification

The reference genome was masked using a custom TE library (Mérel et al. 2021). The Python script create-reads-for-te-sequences.py was used to generate reads corresponding to the TE library using the following parameters: *–read-length 125, –max-error-rate 0, –boost 10* (Kofler et al. 2016; Paris et al. 2020). The reads were then mapped to the reference genome using *bwa bwasw* (*v. 0.7.17*) (Li and Durbin 2010). Aligned bases were masked using *bedtools*, *bamtobed*, and *bedtools maskfasta* (*v. 2.20.0*) (Quinlan and Hall 2010). This process of read generation and mapping was repeated 200 times. Note that sequences smaller than 500 bp were removed from the TE library. Forward and reverse reads were mapped separately to a fasta file containing the masked reference genome and the TE library. The mapping was done using bwa bwasw. For each line, the resulting single-end read alignments files were merged using *PoPoolationTE2 se2pe* (*v. 1.10.04*) (Kofler et al. 2016). *PoPoolationTE2* pipeline was used to estimate TE frequencies in each sample (supplementary table S18, Supplementary Material online). The following options were used in the analysis: *–map-quality 15* (*ppileup* module), *–mode joint*, *–signature-window minimumSampleMedian*, *min-valley minimumSampleMedian, –min-count 2* (identify signature module), *–max-otherte-count 2*, *–max-structvar-count 2* (filterSignatures module), *–min-distance -200*, *–max-distance 300* (*pairupSignatures* module). In the *PoPoolationTE2*, hierarchy file was a file allowing multiple slightly diverged sequences to be assigned to one family, and all sequences with cross mapping reads were regrouped in the same family. The cross mapping was investigated by generating TE reads using create-reads-for-te-sequences.py (*–read-length 125, –max-error-rate 0, –boost 50*) and mapping the reads to the TE library using *bwa bwasw*.

The software was run using the S29, W120, and MT47 DNAseq data. Using the gene annotation of the reference *D. suzukii* genome we identified TE insertions present in genes (exon, intron, 5′ and 3′ UTR) and ±2 kb flanking regions. We consider genes and the ±2 kb flanking regions as having a TEs when the insertion frequency was superior or equal to 0.5 (supplementary table S18, Supplementary Material online). We used a chi-square test to check for the equal distribution of TE in the genomes. We then tested if the presence of TE insertions was independent of the state of the genes (DE or not) using a chi-square test.

### TE Expression Analysis

TE expression was quantified using the *TEcount* module from the *TEtools* software (Lerat et al. 2017). Briefly, *TEcount* sums reads aligned against copies of each TE family annotated from the reference genome creating an output table of expression arranged by TE family (Paris et al. 2020), (supplementary table S19, Supplementary Material online). Differential expression of TEs between paraquat-treated and control flies for each isofemale line was computed using a merged file with the RNA counts for genes and TE families, and following normalization using *DESeq2*.

### TFBM Screening

TE sequences inserted in flanking regions located ±2 kb from DE genes were screened for TFBM (supplementary tables S20 and S21, Supplementary Material online). We selected three TFBM (CnC, Hsf, and dl) related to ARE family from the literature (Villanueva-Cañas et al. 2019). TFBM were screened in *R* (*v. 3.6.0*) using the *JASPAR2020* database R library (*v. 0.99.1*) and *TFBSTools R* library (*v. 1.22.0*) (Tan and Lenhard 2016; Fornes et al. 2020). Position frequency matrix matrices were extracted (CnC: MA0530.1, HSsf: MA1458.1, dl: MA0022.1) before a position weight matrix conversion with the *pseudocount* value set to 0.8. The minimum score value for the screening was fixed at 0.95 to minimize false positives due to small TFBM sequence sizes. P-values were adjusted with the Benjamini–Hochberg correction for multitesting (Benjamini and Hochberg 1995).

## Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

## Acknowledgments

Experimental procedures were supported by the ANR (grant SWING ANR to P.G. and C.V. & grant ExHyb ANR to C.V.) and the Rovaltain Foundation (EpiRip project). Bioinformatic work was performed using the computing facilities of the CC LBBE/PRABI. We thank S. Vicaire for the RNA libraries preparation, K. McKean for English editing and suggestions in the manuscript, N. Burlet, S. Martinez and H. Henri for various technical contributions.

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

## Author Contributions

P.M. produced data, conceived, and wrote the manuscript draft. C.V. and P.G. designed experiments, edited the manuscript. J.P. trimmed NGS data and help in bioinformatic analysis. A.J. calibrated experimental design and produced data. M.L.D.M. supervised the statistical analysis and revised manuscript. M.F. helped on transcriptomic analysis and manuscript correction. M.G.F. performed gene ontology analysis. V.M. produced all related TE information (genome annotations, frequency insertions). All authors proofread and approved the manuscript.

## Data Availability

DNA and RNA-seq data are available on NCBI with the Sequence Read Archive (SRA) number: SUB7319267 and SUB7021358.

## Literature Cited


Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215(3):403–410.

Asplen MK, et al. 2015. Invasion biology of spotted wing Drosophila (*Drosophila suzukii*): a global perspective and future priorities. J Pest Sci. 88(3):469–494.

Austad SN, Fischer KE. 2016. Sex differences in lifespan. Cell Metab. 23(6):1022–1033.

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. J Stat Soft. 67(1):1–48.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc Ser B (Methodol). 57(1):289–300.

Bischoff V, et al. 2006. Downregulation of the Drosophila immune response by peptidoglycan-recognition proteins SC1 and SC2. PLOS Pathog. 2(2):e14.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120.

Bradshaw AD. 1965. Evolutionary significance of phenotypic plasticity in plants. In: Caspari EW and Thoday JM, editors. Advances in genetics. Vol. 13. Amsterdam: Academic Press. p. 115–155.

CABI. Invasive species compendium. Available from: https://www.cabi.org/ISC (accessed July 21, 2021).

Calabria G, Máca J, Bächli G, Serra L, Pascual M. 2012. First records of the potential pest species *Drosophila suzukii* (Diptera: Drosophilidae) in Europe. J Appl Entomol. 136(1 – 2):139–147.

Casacuberta E, González J. 2013. The impact of transposable elements in environmental adaptation. Mol Ecol. 22(6):1503–1517.

Chabrerie O, et al. 2019. Biological invasion theories: merging perspectives from population. Commun Ecosyst Scales. doi:10.20944/preprints201910.0327.v1.

Chen J, Nolte V, Schlötterer C. 2015. Temperature stress mediates decanalization and dominance of gene expression in *Drosophila melanogaster*. PLOS Genet. 11(2):e1004883.

Choudhury RR, Parisod C. 2017. Jumping genes: genomic ballast or powerhouse of biological diversification. Mol Ecol. 26(18):4587–4590.

Czech A, Wende S, Mörl M, Pan T, Ignatova Z. 2013. Reversible and rapid transfer-RNA deactivation as a mechanism of translational repression in stress. PLOS Genet. 9(8):e1003767.

Dalton DT, et al. 2011. Laboratory survival of *Drosophila suzukii* under simulated winter conditions of the Pacific Northwest and seasonal field trapping in five primary regions of small and stone fruit production in the United States. Pest Manag Sci. 67(11):1368–1374.

Deepashree S, Niveditha S, Shivanandappa T, Ramesh SR. 2019. Oxidative stress resistance as a factor in aging: evidence from an extended longevity phenotype of *Drosophila melanogaster*. Biogerontology 20(4):497–513.

Delignette-Muller ML, Dutang C. 2015. fitdistrplus: an R package for fitting distributions. J Stat Soft. 64(4):1–34.

Dubin MJ, Mittelsten Scheid O, Becker C. 2018. Transposons: a blessing curse. Curr Opin Plant Biol. 42:23–29.

Fablet M, Vieira C. 2011. Evolvability, epigenetics and transposable elements. Biomol Concepts 2(5):333–341.

Fornes O, et al. 2020. JASPAR 2020: update of the open-access database of transcription factor binding profiles. Nucleic Acids Res. 48(D1):D87–D92.

Fraimout A, et al. 2017. Deciphering the routes of invasion of *Drosophila suzukii* by means of ABC random forest. Mol Biol Evol. 34(4):980–996.

Ghalambor CK, McKAY JK, Carroll SP, Reznick DN. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. Funct Ecol. 21(3):394–407.

Gibson G. 2008. The environmental contribution to gene expression profiles. Nat Rev Genet. 9(8):575–581.

Girardot F, Monnier V, Tricoire H. 2004. Genome wide analysis of common and specific stress responses in adult *Drosophila melanogaster*. BMC Genomics 5(1):74.

Guerreiro MPG. 2012. What makes transposable elements move in the Drosophila genome? Heredity (Edinb). 108(5):461–468.

Guio L, Barrón MG, González J. 2014. The transposable element Bari-Jheh mediates oxidative stress response in Drosophila. Mol Ecol. 23(8):2020–2030.

Guio L, Vieira C, González J. 2018. Stress affects the epigenetic marks added by natural transposable element insertions in *Drosophila melanogaster*. Sci Rep. 8(1):12197.

Horváth V, Merenciano M, González J. 2017. Revisiting the relationship between transposable elements and the eukaryotic stress response. Trends Genet. 33(11):832–841.

Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. 2010. Inferring regulatory networks from expression data using tree-based methods. Plos One. 5(9):e12776.

Janga SC, Contreras-Moreira B. 2010. Dissecting the expression patterns of transcription factors across conditions using an integrated network-based approach. Nucleic Acids Res. 38(20):6841–6856.

Jangam D, Feschotte C, Betrán E. 2017. Transposable element domestication as an adaptation to evolutionary conflicts. Trends Genet. 33(11):817–831.

Kaminker JS, et al. 2002. The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. Genome Biol. 3(12):RESEARCH0084.

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 37(8):907–915.

King AM, MacRae TH. 2015. Insect heat shock proteins during stress and diapause. Annu Rev Entomol. 60:59–75.

Kofler R, Gómez-Sánchez D, Schlötterer C. 2016. PoPoolationTE2: comparative population genomics of transposable elements using Pool-Seq. Mol Biol Evol. 33(10):2759–2764.

Kohen R, Chevion M, Czapski G. 1985. Transition metals potentiate paraquat toxicity. Free Radic Res Commun. 1(2):79–88.

Kuznetsova A, Brockhoff PB, Christensen RHB. 2017. lmerTest Package: tests in linear mixed effects models. J Stat Soft. 82(13):1–26.

Lambert SA, et al. 2018. The human transcription factors. Cell 172(4):650–665.

Langille AB, Arteca EM, Newman JA. 2017. The impacts of climate change on the abundance and distribution of the Spotted Wing Drosophila (*Drosophila suzukii*) in the United States and Canada. PeerJ. 5:e3192.


Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021

Lasky JR, Forester BR, Reimherr M. 2018. Coherent synthesis of genomic associations with phenotypes and home environments. Mol Ecol Resour. 18(1):91–106.
Lavrinienko A, et al. 2017. First record of the invasive pest *Drosophila suzukii* in Ukraine indicates multiple sources of invasion. J Pest Sci. 90(2):421–429.
Lerat E, Fablet M, Modolo L, Lopez-Maestre H, Vieira C. 2017. TEtools facilitates big data expression analysis of transposable elements and reveals an antagonism between their activity and that of piRNA genes. Nucleic Acids Res. 45(4):e17.
Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 26(5):589–595.
Lin Y, Chen Z-X, Oliver B, Harbison ST. 2016. Microenvironmental gene expression plasticity among individual *Drosophila melanogaster*. G3 (Bethesda) 6(12):4197–4210.
Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15(12):550.
Magwire MM, Bayer F, Webster CL, Cao C, Jiggins FM. 2011. Successive increases in the resistance of drosophila to viral infection through a transposon insertion followed by a duplication. PLOS Genet. 7(10):e1002337.
Marin P, et al. 2020. Biological invasion: the influence of the hidden side of the (epi)genome. Funct Ecol. 34(2):385–400.
McClintock B. 1950. The origin and behavior of mutable loci in Maize. Proc Natl Acad Sci U S A. 36(6):344–355.
Mérel V, et al. 2021. The worldwide invasion of *Drosophila suzukii* is accompanied by a large increase of transposable element load and a small number of putatively adaptive insertions. Mol Biol Evol. doi:10.1093/molbev/msab155.
Modolo L, Lerat E. 2015. UrQt: an efficient software for the unsupervised Quality trimming of NGS data. BMC Bioinformatics 16:137.
Moran EV, Alexander JM. 2014. Evolutionary responses to global change: lessons from invasive species. Ecol Lett. 17(5):637–649.
Morrow G, Le Pécheur M, Tanguay RM. 2016. *Drosophila melanogaster* mitochondrial Hsp22: a role in resistance to oxidative stress, aging and the mitochondrial unfolding protein response. Biogerontology 17(1):61–70.
Moskalev AA, et al. 2019. Transcriptome analysis of long-lived *Drosophila melanogaster* E (z) mutants sheds light on the molecular mechanisms of longevity. Sci Rep. 9(1):9151.
Naito K, et al. 2009. Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. Nature 461(7267):1130–1134.
Olazcuaga L, et al. 2020. A whole-genome scan for association with invasion success in the fruit fly *Drosophila suzukii* using contrasts of allele frequencies corrected for population structure. Mol Biol Evol. 37(8):2369–2385.
De Oliveira MV, et al. 2013. Gene expression modulation by paraquat-induced oxidative stress conditions in *Paracoccidioides brasiliensis*. Fungal Genet Biol. 60:101–109.
Paris M, et al. 2020. Near-chromosome level genome assembly of the fruit pest *Drosophila suzukii* using long-read sequencing. Sci Rep. 10(1):11227.
Pasyukova EG, Roshina NV, Mackay TFC. 2004. Shuttle craft: a candidate quantitative trait gene for *Drosophila* lifespan. Aging Cell. 3(5):297–307.
Peignier S, Schmitt P, Calevro F. 2020. GReNaDIne: data-driven approaches to infer gene regulatory networks in python. https://hal.archives-ouvertes.fr/hal-02863880 (accessed July 25, 2021).
Pokhilko A, et al. 2021. Targeted single-cell RNA sequencing of transcription factors enhances the identification of cell types and trajectories. Genome Res. 31(6):1069–1081.
Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26(6):841–842.
Rebollo R, Romanish MT, Mager DL. 2012. Transposable elements: an abundant and natural source of regulatory sequences for host genes. Annu Rev Genet. 46:21–42.
Rey O, Danchin E, Mirouze M, Loot C, Blanchet S. 2016. Adaptation to global change: a transposable element-epigenetics perspective. Trends Ecol Evol. 31(7):514–526.
Roberts A, Pachter L. 2013. Streaming fragment assignment for real-time analysis of sequencing experiments. Nat Methods 10(1):71–73.
Rzezniczak TZ, Douglas LA, Watterson JH, Merritt TJS. 2011. Paraquat administration in Drosophila for use in metabolic studies of oxidative stress. Anal Biochem. 419(2):345–347.
Sage RF. 2020. Global change biology: a primer. Glob Chang Biol. 26(1):3–30.
Schrader L, et al. 2014. Transposable element islands facilitate adaptation to novel environments in an invasive species. Nat Commun. 5:5495.
Shannon P, et al. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13(11):2498–2504.
Shaposhnikov M, Proshkina E, Shilova L, Zhavoronkov A, Moskalev A. 2015. Lifespan and stress resistance in Drosophila with overexpressed DNA repair genes. Sci Rep. 5:15299.
Slotkin RK, Martienssen R. 2007. Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet. 8(4):272–285.
Song Y, et al. 2015. Chilling- and freezing- induced alterations in cytosine methylation and its association with the cold tolerance of an alpine subnival plant, *Chorispora bungeana*. Plos One 10(8):e0135485.
Spannhoff A, et al. 2011. Histone deacetylase inhibitor activity in royal jelly might facilitate caste switching in bees. EMBO Rep. 12(3):238–243.
Srikant T, Drost H-G. 2020. How stress facilitates phenotypic innovation through epigenetic diversity. Front Plant Sci. 11:606800.
Stapley J, Santure AW, Dennis SR. 2015. Transposable elements as agents of rapid adaptation may explain the genetic paradox of invasive species. Mol Ecol. 24(9):2241–2252.
Stephens M. 2017. False discovery rates: a new deal. Biostatistics 18(2):275–294.
Supek F, Bošnjak M, Škunca N, Šmuc T. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. PLOS One 6(7):e21800.
Tan G, Lenhard B. 2016. TFBSTools: an R/bioconductor package for transcription factor binding site analysis. Bioinformatics 32(10):1555–1556.
Theoharides KA, Dukes JS. 2007. Plant invasion across space and time: factors affecting nonindigenous species success during four stages of invasion. New Phytol. 176(2):256–273.
Tsai W-T. 2013. A review on environmental exposure and health risks of herbicide paraquat. Toxicol Environ Chem. 95(2):197–206.
Tsai W-T. 2020. Status of herbicide use, regulatory management and case study of paraquat in Taiwan. Environ Dev Sustain. 22(3):2673–2683.
Ullastres A, Merenciano M, Guio L, González J. 2016. Transposable elements: a toolkit for stress and environmental adaptation in bacteria. In: Stress and environmental regulation of gene expression and adaptation in bacteria. New Jersey: John Wiley & Sons, Ltd. p 137–145.
Villanueva-Cañas JL, Horvath V, Aguilera L, González J. 2019. Diverse families of transposable elements affect the transcriptional regulation of stress-response genes in *Drosophila melanogaster*. Nucleic Acids Res. 47(13):6842–6857.
Weber AL, et al. 2012. Genome-wide association analysis of oxidative stress resistance in *Drosophila melanogaster*. PLOS One 7(4):e34745.
Zer H, Freedman JH, Peisach J, Chevion M. 1991. Inverse correlation between resistance towards copper and towards the redox-cycling compound paraquat: a study in copper-tolerant hepatocytes in tissue culture. Free Radic Biol Med. 11(1):9–16.
Zou S, Meadows S, Sharp L, Jan LY, Jan YN. 2000. Genome-wide study of aging and oxidative stress response in *Drosophila melanogaster*. Proc Natl Acad Sci U S A. 97(25):13726–13731.

**Associate editor:** Sarah Schaack

Downloaded from https://academic.oup.com/gbe/article/13/9/evab208/6368064 by guest on 21 October 2021