



HAL
open science

Drosophila suzukii oxidative stress response involves Jheh gene cluster but not transposable elements

Pierre Marin, Angelo Jacquet, H el ene Henri, Patricia Gibert, Cristina Vieira

► **To cite this version:**

Pierre Marin, Angelo Jacquet, H el ene Henri, Patricia Gibert, Cristina Vieira. Drosophila suzukii oxidative stress response involves Jheh gene cluster but not transposable elements. 2020. hal-03014540

HAL Id: hal-03014540

<https://cnrs.hal.science/hal-03014540v1>

Preprint submitted on 19 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destin ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.

1 **Title: *Drosophila suzukii* oxidative stress response involves *Jheh* gene cluster but not transposable**
2 **elements**

3

4 **Running title: *D. suzukii* stress response and *Jheh***

5

6 Authors : Pierre Marin¹, Angelo Jacquet¹, H el ene Henri¹, Patricia Gibert¹ & Cristina Vieira¹

7 ¹Laboratoire de Biom etrie et Biologie Evolutive UMR 5558, CNRS, Universit  Lyon 1, Universit 
8 de Lyon, Villeurbanne, France

9 Corresponding author's email : patricia.gibert@univ-lyon1.fr, cristina.vieira@univ-lyon1.fr

10

11 **Keywords:** invasive species, gene expression, rapid adaptation, paraquat, genotypes

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30 **Summary statement**

31 The responses to oxidative stress of the invasive species, *Drosophila suzukii*, show variability
32 between genotypes related to their invasion status. The genes of the juvenile hormone epoxide
33 hydrolase cluster are involved in this response.

34

35 **ABSTRACT**

36 The study of the mechanisms involved in adaptation remains a timely issue, particularly in the
37 context of global changes. To better understand these mechanisms of rapid adaptation, invasive
38 species are a good model because they are subjected to new and/or different environmental
39 factors. Using different lines of different geographical origin of the invasive pest *Drosophila*
40 *suzukii*, we characterized the phenotypic response to oxidative stress. Subsequently, we tested
41 the involvement of the *Jheh* gene cluster in this response and the possible role of transposable
42 elements. We show that the resistance to oxidative stress of the lines appears to be related to
43 their invasive status and we confirm the role of the *Jheh* gene cluster in this response. We have
44 not identified any transposable elements in this gene region that could influence the expression
45 of the gene.

46

47 **INTRODUCTION**

48 The rapid spread of invasive species in a huge spectrum of environments relies on multiple
49 factors, from genetics to phenotypic plasticity, probably including fine molecular mechanism such
50 as hormonal production or epigenetic gene regulation (Beldade *et al.*, 2011; Marin *et al.*, 2019;
51 Stapley *et al.*, 2015). Phenotypic plasticity, i.e., the ability of a genotype to express different
52 phenotypes in different environments (Ghalambor *et al.*, 2015) has been proposed as one of the
53 most promising explanations for invasive success, particularly in the case of founder population
54 depleted of genetic variation (Estoup *et al.*, 2016; Marin *et al.*, 2019). Among deleterious
55 environments that can be encountered by invasive species, oxidative stress caused by
56 phytosanitary products is one of them. The invasive pest, *Drosophila suzukii*, is a good model to
57 investigate the adaptive process during invasion (Gibert *et al.*, 2016). This species which belong
58 to the group of the fruit fly *D. melanogaster*, originally comes from Asia and was detected
59 simultaneously both in North America (U.S.A) and in Europe in 2008. North America was invaded
60 by native Japan populations derived from Hawaii. In Europe, several introductions were detected
61 from U.S.A and from China (Frainout *et al.*, 2017). Currently, *D. suzukii* is present in both North
62 and South America, in Europe from the south (Spain) to the East (Poland, Ukraine) and it has
63 also been observed in Russia (CABI, 2020; Lavrinienko *et al.*, 2017).

64 Characterization of the phenotypic and molecular responses of *D. suzukii* to changing
65 environmental conditions may provide information to the mechanisms involved in the ability of
66 invasive species to cope with environmental variation. Paraquat (N,N'-dimethyl-4,4'-bipyridinium
67 dichloride) is one of the most widely used herbicide in the world leading to the production of ROS
68 (reactive oxygen species) (Tsai, 2018). Oxidative stress due to the use of paraquat in the field
69 has also been used in the laboratory as a good proxy for studying stress resistance (Bus J S and
70 Gibson J E, 1984; Rzezniczak *et al.*, 2011). Paraquat was banned since 2007 in Europe but is
71 still used in many other regions like in U.S.A or Japan. Paraquat exposition is known to induce a
72 reduction in the lifespan associated with changes in gene expression (Finkel and Holbrook, 2000;
73 Liguori *et al.*, 2018; Vermeulen *et al.*, 2005). One of the candidate genes involved in paraquat
74 resistance is the cluster of *Jheh* (*Juvenile hormone epoxide hydrolase*) genes, which are not only
75 involved in the lifespan but also in response to the oxidative environment (Flatt and Kawecki,
76 2007; Guio *et al.*, 2014). Moreover in *D. melanogaster*, an insertion of a transposable element
77 (TE) *Bari-Jheh*, near the cluster of the *Jheh* genes has been described as driving an increase of
78 resistance in presence of paraquat (Guio *et al.*, 2014).

79 Using several strains of *D. suzukii*, we measured responses to oxidative stress at the phenotypic
80 and molecular level. We made the hypothesis that different genetic backgrounds from native and
81 invasive populations will have different responses to oxidative stress and that the *Jheh* cluster
82 may be involved on it. Due to the over-representation of TEs in the genome of *D. suzukii* (33% of
83 the repeated elements, (Sessegolo *et al.*, 2016)), compared to other *Drosophila* species, we
84 looked for the presence of TEs in this region in the different lines. We monitored lifespan after
85 paraquat exposure and measured the expression of three genes of *Jheh* cluster *Jheh-1*, *Jheh-2*
86 and *Jheh-3* in six isofemale lines, four from the invasive regions, North America (Watsonville and
87 Dayton) and France (Paris and Montpellier) and two from the native area, Japan (Sapporo and
88 Tokyo). We evaluated the genetic diversity within and between lines by sequencing introns of the
89 *Jheh* genes, searched for TEs and for transcription factor binding sites (TFBS). Our results
90 suggest a strong effect of the genotype on the resistance to stress and changes in *Jheh*
91 expression levels, with no link with TEs.

92

93 MATERIAL AND METHODS

94 *Drosophila suzukii* lines and rearing conditions

95

96 *D. suzukii* lines were sampled in 2014 from one native country (Japan: Sapporo and Tokyo) and
97 2 invaded areas (USA: Watsonville and Dayton and France: Montpellier and Paris). Field-
98 inseminated females were isolated to establish half-sib families called isofemale lines commonly

99 used to investigate *Drosophila* natural populations (David *et al.*, 2005). Flies were reared in
100 modified medium (drosophila agar type, ref.66-103, Apex™, 9 g.L⁻¹ ; cornmeal 33 g.L⁻¹ ; yeast,
101 dried yeast, ref.75570, LYNSIDE® 17 g.L⁻¹ ; industrial sugar 50 g.L⁻¹ ; nipagin, Tegosept, ref.20-
102 258, Apex™ 4 g.L⁻¹ ; 96% ethanol 40 ml.L⁻¹ ; distilled water 1 L) from Dalton *et al.*, (2011), in a
103 humidified, temperature-controlled incubator at 22.5°C, 70 % of relative humidity and a 16:8 LD
104 cycle. The recipe of the modified medium was to bring to boil agar, cornmeal, yeast extract and
105 sugar in distilled water. Then wait out of the fire about 10 minutes until the mixture cooled to 53°C
106 before adding diluted nipagin in 96% ethanol. Medium is then poured in vials and cooled at room
107 temperature before to be stored at 4°C. All the experiments were made with 4 to 7 days old flies.

108 **Oxidative stress resistance experiments**

109 We used paraquat (methyl viologen dichloride hydrate, ref. 75365-73-0, Sigma-Aldrich®) to mimic
110 oxidative stress. Oxidative stress was assessed by adding paraquat directly in the medium (10
111 mM) before the cooling step and below 53°C. The control experiment was made with the same
112 medium but without paraquat. We used one isofemale line per locality (total of six) named
113 Montpellier (France), Paris (France), Sapporo (Japan), Tokyo (Japan), Dayton (U.S.A.) and
114 Watsonville (U.S.A.). We made three replicates per line and per sex, with ten flies per replicate.
115 Survival was monitored every 24h. Flies were transferred into new vials every three to four days
116 to limit microbial development.

117 **RT-qPCR analysis of *Jheh* genes**

118 We quantified the expression of the three *Jheh* genes (*Jheh-1*, *Jheh-2* and *Jheh-3*) by RT-qPCR
119 after induction of oxidative stress and in control condition. Adult males and females were exposed
120 during 24h to medium culture with 20 mM of paraquat. After 24h the flies were immediately
121 dissected in PBS 1X solution (Gibco Thermo-Fisher) in order to extract carcasses for both sexes
122 and eliminate germline tissues. We made three replicates per sex and treatment and used four
123 flies per replicate.

124 RNA extraction was made using Direct-zol™-96 RNA Kits (Zymo Research), following the
125 manufacturer recommendations and RNA was treated with DNase. cDNA were obtained from 0.5
126 µg of RNA using SuperScript™ IV VIL0™ Master Mix (Invitrogen). RT negative control was made
127 with RNA but without the reverse transcriptase to control for genomic DNA contamination. cDNA
128 were stored at -80°C before the quantification step. Gene expression was then quantified by
129 quantitative PCR and *Rp49* was used as housekeeping gene. Primers were designed using the
130 *D. suzukii* referenced genome (Table S1, (Chiu *et al.*, 2013)). Their efficiency was between 91.1%
131 to 97.2% (*RP49*: 91.6 %, *Jheh-1*: 97.2%, *Jheh-2*: 95.2%, *Jheh-3*: 91.1%). 2 µl of the cDNA sample
132 were supplemented with 5 µL of SsoADV Universal SYBR Green Supermix (BioRad) mix 2X, 0.3

33 μl of each primer (10 μM) and 2.4 μl of pure water. We made technical duplicates for each sample.
34 PCR reactions were made in a BioRAAd CFX-96 with a program consisted of an initial activation
35 of 95°C for 10 minutes and then 40 cycles each comprising 15 seconds at 95°C, 10 seconds at
36 60°C and 72°C .

37 **Genetic diversity of isofemale lines**

38 We sequenced intronic regions of *Jheh* gene cluster of the six lines used in this study. DNA was
39 extracted individually from 10 females per line with the 96-Well Plate Animal Genomic DNA
40 Miniprep kit (ref. BS437, Biobasic) following the manufacturer instructions. Primers were designed
41 to flank the intronic regions for the three *Jheh* genes (Table S1) and Phusion high fidelity DNA
42 Polymerase (2 U/ μL) (F-530XL Thermofisher Scientific) was used to amplify sequences. The
43 same PCR program was used for all primers pairs: 98°C for 10 minutes, followed by 40 cycles
44 composed of 30 seconds at 98°C, 1 minute at 56°C and 20 seconds at 72°C and a final elongation
45 step for 1 minute at 72°C. The sequencing of the two strands was done directly from the PCR
46 product by BIOFIDAL sequencing company (Vaulx en Velin, France). Sequences were manually
47 curated with CLC Main Workbench 8 software (Qiagen) before being aligned with the Muscle
48 program implemented in the workbench to generate haplotypes by line for each intron. MEGA X
49 software was used to calculate pairwise comparison and nucleotide diversity using p-method
50 option (Table S2-S3) (Kumar *et al.*, 2018).

51 **Detection of Transposable elements and transcription factor binding sites**

52 We sequenced the intergenic regions of the *Jheh* gene cluster, plus the 5' and 3' regions
53 of the cluster (Table S1). DNA was extracted from one female per population as described above.
54 Classical PCR method was used with the following program, 10 minutes at 95°C followed by
55 several cycles composed of 30 seconds at 95°C, 30 seconds at 63°C, 3 minutes at 72°C and a
56 final elongation of 15 minutes at 72°C. The number of cycles was 25 for the region before *Jheh-1*
57 and between *Jheh-1* and *Jheh-2*, 35 cycles for the region between *Jheh-2* and *Jheh-3* and 30
58 cycles after *Jheh-3*. We identified TEs in the intergenic regions by a blast against a homemade
59 data base of the TE sequences from the *D. suzukii* reference genome (Paris *et al.*, 2020, Mérel
60 *et al.*, *in prep.*).

61 For TFBS (Table 1), we used conSite website to screen all TFBS from insect in our sequences
62 (Sandelin *et al.*, 2004). To complete our analysis, we used the TFBS obtained from Villanueva-
63 Cañas *et al.* (2019) and we extracted PFM (position frequency matrix) of the 14 TFBS from the
64 JASPAR2018 database (v.1.1.1) (Parcy *et al.*, 2017). Then, we used TFBSTools (v.1.22.0)
65 package from R software (v. 3.6.0) to convert in PWM (position weight matrix), and then search

66 on the 6 lines and the reference genome of *D. suzukii* (Paris *et al.*, 2020; R Core Team, 2019;
67 Tan and Lenhard, 2016).

68 **Statistical analysis**

69 **Survival monitoring**

70 Survival data were analyzed using a linear mixed model with lmer function from lme4 provided on
71 R (v. 3.6.0) after a log transformation, confirmation of normality and homoscedasticity (Bates *et*
72 *al.*, 2015). This model was chosen after log-likelihood comparison between models (linear model
73 with raw or log transformed data, survival model with a Weibull distribution).

74 We analyzed sexes separately to limit interaction terms, and focused on the effect of the
75 treatment, the lines and their interaction. Biological replicates were added as random effect and
76 we plotted exponential of the values and associated confidence interval on the Fig. S1. Those
77 effects can be interpreted as multiplicative effect on the mean lifespan compared to the reference
78 chosen here as the non-exposed group from Sapporo (*e.g.* the Sapporo reference is centered on
79 1 and the effect of paraquat 0.18 involves a survival time under paraquat for Sapporo of 0.18 or
80 18% of the survival time without paraquat).

81

82 **qPCR analysis**

83 RT-qPCR raw data were analyzed using R and EasyqpcR library (1.21.0) for the quantification
84 and normalization with *RP49* (Sylvain, 2012). Data were analyzed separately for the three genes
85 (*Jheh-1 -2* and *-3*) and sex using a linear model (ANOVA2, Table S4) after log transformation to
86 validate homoscedasticity and normality. Pairwise comparisons were made using a Tuckey test.

87 **RESULTS**

88 ***D. suzukii* wild type lines have significant differences in life span**

89 To investigate the influence of the genotypes from different geographical origins on the
90 lifespan, we compared the invasive and native *D. suzukii* lines in control condition (Fig. 1A, Fig.
91 S1 & Table 2). The lifespan ranges from 31 to 55 days for females and from 25 to 45 days for
92 males. For females we observed a strong genotype effect related to geographic location: the
93 genotypes that lived the longest were those of Dayton and Paris (about 1.88-1.96 times more
94 than Sapporo, Fig. S1). Sapporo, Tokyo, and Watsonville were not significantly different and with
95 the lowest lifespan. For males, the four invasive genotypes from Europe and U.S.A had a higher
96 lifespan than Sapporo. Tokyo was similar to Sapporo.

97 As expected, exposure to paraquat reduced life span on average from 82 to 77% for
98 females and males (Fig. 1A, Fig. S1). The two lines with the best paraquat resistance in absolute
99 value were still Dayton and Paris in both sexes (Fig. 1A and Table 2). We then wanted to have
00 an estimate of paraquat sensitivity (i.e., the slope difference Fig.1 B) taking into account the
01 longevity of each line by estimating the value of the interaction coefficients (i.e., the slope
02 difference compared to Sapporo) in Table 2 and statistically tested in Fig. S1. Again, the effect
03 was not similar between genotypes and sexes. For females, Paris was the line presenting
04 significantly the highest sensitivity (-0.87, Table 2) with a reduction of 28% of the life span
05 compared to Sapporo (Fig. S1). For males, the reduction in life span was significantly the highest
06 for Montpellier and Watsonville (-0.84 and -0.83) with a reduction from 34 and 32% by comparison
07 with Sapporo. These results reveal a strong genotype-by-environment interaction in the response
08 to oxidative stress and also a sex effect. It is interesting to note that despite the shorter life span
09 of Japanese genotypes, and in particular of Sapporo in the absence of treatment, these genotypes
10 were the most resistant to paraquat exposure, as shown by the lowest ratio of paraquat lifetime
11 to control lifetime (Table 2).

12

13 ***Jheh* genes expression changes with the paraquat treatment**

14 To investigate the effect of paraquat-mediated oxidative stress on the gene expression
15 level, we focused on *Jheh* gene cluster described as potentially involved in stress response in
16 insects and mammals (Guio *et al.*, 2014; Oesch *et al.*, 2000). We quantified the level of expression
17 of the *Jheh* genes (*Jheh-1*, *Jheh-2* and *Jheh-3*) in adult males and females flies for the six
18 genotypes described above (Fig. 2).

19 We observed strong differences between males and females. For males, gene expression
20 was not significantly different between control and paraquat treatment for the six genotypes and
21 for the three gene. In females, the effect of paraquat was different according to the gene and the
22 genotype (Table S4). For *Jheh-1* and *Jheh-2*, oxidative stress resulted in a significant increase of
23 gene expression for the two French genotypes and the Tokyo genotype. On the contrary, the
24 Sapporo genotype exhibited a significant reduction of *Jheh-1* expression in presence of paraquat.
25 For *Jheh-3*, we observed a downregulation of the gene expression only for the Sapporo genotype.

26 **Low Genetic diversity of lines in *Jheh* cluster**

27 To assess the levels of neutral genetic diversity within and between lines, we sequenced intronic
28 regions for *Jheh* genes for each genotype (Fig. 3). As expected, the within-line polymorphism was
29 very low (Table 3, Fig. S2), with the exception of Watsonville with 0.0792 for the first intron of
30 *Jheh-1*. The number of haplotypes was also low (Table S2). The first intron of *Jheh-2* presents

31 the highest levels of diversity, contrasting with the other introns. This corresponds to a residual
32 polymorphism that is still present in the lines despite the laboratory rearing.

33 Depending on the intronic regions we found between two to four haplotypes per genotype
34 (Table S2). We computed the diversity between genotypes (global intronic nucleotide diversity π)
35 using the most common haplotype for each intron, showing that on average these values are very
36 small, with the highest value for *Jheh-2.1* as mentioned above (Table S3).

37 **Jheh harbour transcription factor binding sites (TFBS)**

38
39 Transcription factor binding site (TFBS) are *cis* regulatory sequences that are recognized
40 by transcription factors and modify gene expression. Several TFBS are known to be involved
41 during oxidative response. We detected 9 of the 14 previous identified TFBS in the intergenic
42 regions: HIF1A, br, cad, Cf2, Deaf1, CnC, dl, hb and Ubx (Fig. 3 & 4, Table S5). Comparison of
43 the number of TFBS between genotypes (Fig. 4) revealed several differences but not clear link
44 with the changes in expression observed for the *Jheh* genes. For example, the Sapporo genotype
45 which consistently showed a decrease in the expression of all three genes, did not appear to have
46 a different specific TFBS. The two French genotypes which exhibited systematically an increase
47 of expression after paraquat treatment appeared to have an increased number of putative TFBS.
48 For example, the two French genotypes showed two Deaf1 motives when compared to the other
49 genotypes upstream of TSS of the *Jheh-1* gene. In the case of the *Jheh-2* gene, the French
50 genotypes presented a significant number of TFBS, with for example six TFBS for the Montpellier
51 genotype (2Ubx, 2hb, CnC and cad). In this region, no genotype showed the same pattern of
52 TFBS and it was similar for *Jheh-3*.

53

54 **Transposable elements do not affect *Jheh* gene expression**

55 The presence of TE in the vicinity or within the *Jheh* cluster could impact the gene
56 expression during oxidative stress because they could bring Antioxidant Response Element for
57 transcription factors or by modifying chromatin state (Guio and González, 2015; Guio *et al.*, 2014).
58 Surprisingly, and even if *D. sukuzii* harbors more than 30% of TEs, no full insertion was observed
59 in the *Jheh* cluster indicating that we are probably in regions of high recombination. However, we
60 did identify small pieces of TE that are quite conserved between the genotypes but no TFBS were
61 detected in these sequences (Fig. 3, Table 4 & S5). No obvious link seems to exist between gene
62 expression and the presence of TE in the *Jheh* cluster.

63 **DISCUSSION**

64

65 A growing body of literature suggests that responses to oxidative stress in *Drosophila* may
66 be mediated by insertions of TEs, that in some cases could affect gene expression or the
67 chromatin structure (Guio *et al.*, 2014). In *D. melanogaster*, the *Jheh* gene cluster has been shown
68 to be involved in the response to paraquat treatment and associated with local adaptation. Guio
69 *et al.* (2014) compared *D. melanogaster* genotypes with and without Bari-Jheh TE insertion, and
70 showed that, (i) TE insertion had a cost in the absence of stress, (ii) TE insertion confer increased
71 survival in the presence of oxidative stress, (iii) TE insertion provides antioxidant response
72 elements (AREs) that contribute to altered gene expression (Guio and González, 2015; Guio *et*
73 *al.*, 2014). In this study, we analyzed the expression of the *Jheh* gene cluster in several genotypes
74 of *D. sukuzii* to test whether the *Jheh* gene cluster is involved in the oxidative stress response
75 and whether TEs could also be associated with alterations in gene expression.

76 We first measured the life span of flies without treatment. We showed that flies of the
77 Japanese genotypes exhibited the shortest lifespan in both males and females. Surprisingly,
78 these lines showed an increased resistance to oxidative stress. The French lines were more
79 sensitive to paraquat than the American ones, although notable differences were observed
80 between lines from the same continent. The negative association we observed between longevity
81 and paraquat resistance had not been observed in previous work with *D. melanogaster*, in which
82 the opposite association was observed (Finkel and Holbrook, 2000; Liguori *et al.*, 2018). It could
83 be argued that the use of paraquat in Europe has been banned since 2007, which could lead to
84 a loosening of selection on genes related to paraquat resistance, as observed in other organisms
85 (Campos *et al.*, 2014; Shaw, 2000).

86 We then measured the expression of the *Jheh* genes previously reported to be involved in
87 the oxidative response. Consistent with the literature of *D. melanogaster*, we found sex-specific
88 responses to oxidative stress (Guio *et al.*, 2014; Weber *et al.*, 2012). For *Jheh-1* and *Jheh-2*, we
89 observed a significant effect of genotype and treatment, but only for females, contrary to what
90 was reported in *D. melanogaster*. For *Jheh-3*, treatment and genotype effect were significant for
91 both males and females. These differences in gene expression could not be associated with the
92 presence of TEs insertions, since only partial sequences were present in the intergenic regions.
93 The presence of various TFBS could contribute to the observed differences. We also quantified
94 the polymorphism in our lines, which could be associated with differences in gene expression.
95 We did not observe total homozygosity in the lines but genetic diversity was much lower than
96 what is observed in natural populations of *D. melanogaster*. Lack *et al.* (2016) studied populations
97 from several continents and measured values of nucleotide diversity of up to 0.401 within the
98 population. For inbred DGRP (*Drosophila* Genetic Reference Panel) lines, the mean intronic
99 diversity was 0.0076 ± 0.008 , which is close to the values we observed (Mackay *et al.*, 2012). It

00 is therefore unlikely that the residual polymorphism in the *Jheh* gene region can explain the
01 differences in gene expression.

02 The striking result in our analysis is the similar pattern of changes in the expression of
03 *Jheh-1* and *Jheh-2* in females of European genotypes, with an increase in expression, which was
04 associated with lower resistance to oxidative stress, since these are the most sensitive genotypes.
05 On the contrary, the Sapporo genotype systematically showed a reduction in the expression levels
06 of the three genes, which could also be associated with increased resistance in the presence of
07 paraquat, but this was not observed for the Tokyo genotype.

08

09

10 **Conclusion**

11 In conclusion, our work shows for the first time how various genotypes of *D. suzukii*
12 respond to oxidative stress and suggests that populations found in invaded areas are more
13 sensitive than Japanese populations, specially the French ones. We have also confirmed that the
14 *Jheh* gene cluster is involved in the response to oxidative stress also in *D. suzukii*, independently
15 of the presence of TE in intergenic regions. This work also suggests that the genetic background
16 and probably trans regulatory sequences are involved in gene expression and stress response.
17 Further phenotypic and genomic studies on natural populations are needed to better understand
18 the success of invasive species such as *D. suzukii*.

19

20 **Acknowledgements**

21 We thanks to Vincent Mérel who produced the TEs annotation.

22 P. M. produced data, conceived and wrote the manuscript draft. C. V. & P.G. designed
23 experiments, edited the manuscript. A. J. calibrated experimental design. H. H. helped to q-PCR
24 method and analysis.

25 **Funding**

26 Experimental procedures were supported by the ANR (grant SWING ANR to P.G. and C.V. grant
27 ExHyb ANR to C.V.) and the Rovaltain foundation (EpiRip project).

28

29 **Author's informations**

30 Cristina Vieira: <https://orcid.org/0000-0003-3414-3993>

31 Pierre Marin: <https://orcid.org/0000-0002-8304-138X>

32 Patricia Gibert: <https://orcid.org/0000-0002-9461-6820>

33

34 **Competing interests**

35 The authors declare that they have no competing interests.

36

37 **Bates, D., Mächler, M., Bolker, B. M. and Walker, S. C.** (2015). Fitting linear mixed-effects
38 models using lme4. *Journal of Statistical Software* **67**,.

39 **Beldade, P., Mateus, A. R. A. and Keller, R. A.** (2011). Evolution and molecular mechanisms
40 of adaptive developmental plasticity. *Molecular Ecology* **20**, 1347–1363.

41 **Bus J S and Gibson J E** (1984). Paraquat: model for oxidant-initiated toxicity. *Environmental*
42 *Health Perspectives* **55**, 37–46.

43 **CABI** (2020). CABI. Invasive Species Compendium. Wallingford, UK: CAB International. *CABI*.

44 **Campos, M. R., Rodrigues, A. R. S., Silva, W. M., Silva, T. B. M., Silva, V. R. F., Guedes, R.**
45 **N. C. and Siqueira, H. A. A.** (2014). Spinosad and the Tomato Borer *Tuta absoluta*: A
46 Bioinsecticide, an Invasive Pest Threat, and High Insecticide Resistance. *PLoS One* **9**,.

47 **Chiu, J. C., Jiang, X., Zhao, L., Hamm, C. A., Cridland, J. M., Saelao, P., Hamby, K. A., Lee,**
48 **E. K., Kwok, R. S., Zhang, G., et al.** (2013). Genome of *Drosophila suzukii*, the Spotted
49 Wing *Drosophila*. *G3: Genes, Genomes, Genetics* **3**, 2257–2271.

50 **Dalton, D. T., Walton, V. M., Shearer, P. W., Walsh, D. B., Caprile, J. and Isaacs, R.** (2011).
51 Laboratory survival of *Drosophila suzukii* under simulated winter conditions of the Pacific
52 Northwest and seasonal field trapping in five primary regions of small and stone fruit
53 production in the United States. *Pest Management Science* **67**, 1368–1374.

54 **David, J. R., Gibert, P., Legout, H., Pétavy, G., Capy, P. and Moreteau, B.** (2005). Isofemale
55 lines in *Drosophila*: an empirical approach to quantitative trait analysis in natural
56 populations. *Heredity* **94**, 3–12.

57 **Estoup, A., Ravign, V., Hufbauer, R., Vitalis, R., Gautier, M. and Facon, B.** (2016). Is There
58 A Genetic Paradox of Biological Invasion? *Annual Review of Ecology Evolution and*
59 *Systematics* **47**, 51–72.

60 **Finkel, T. and Holbrook, N. J.** (2000). Oxidants, oxidative stress and the biology of ageing.
61 *Nature* **408**, 239–247.

62 **Flatt, T. and Kawecki, T. J.** (2007). Juvenile Hormone as a Regulator of the Trade-Off Between
63 Reproduction and Life Span in *Drosophila Melanogaster*. *Evolution* **61**, 1980–1991.

64 **Fraimout, A., Debat, V., Fellous, S., Hufbauer, R. A., Foucaud, J., Pudlo, P., Marin, J.-M.,**
65 **Price, D. K., Cattel, J., Chen, X., et al.** (2017). Deciphering the routes of invasion of
66 *Drosophila suzukii* by means of ABC random forest. *Molecular biology and evolution* **34**,
67 980–996.

68 **Ghalambor, C. K., Hoke, K. L., Ruell, E. W., Fischer, E. K., Reznick, D. N. and Hughes, K.**
69 **a.** (2015). Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression
70 in nature. *Nature* **525**, 372–375.

71 **Gibert, P., Hill, M., Pascual, M., Plantamp, C., Terblanche, J. S., Yassin, A. and Sgrò, C. M.**
72 (2016). *Drosophila* as models to understand the adaptive process during invasion.
73 *Biological Invasions* **18**, 1089–1103.

- 74 **Guio, L. and González, J.** (2015). The Dominance Effect of the Adaptive Transposable
75 Element Insertion Bari-Jheh Depends on the Genetic Background. *Genome Biol Evol* **7**,
76 1260–1266.
- 77 **Guio, L., Barrón, M. G. and González, J.** (2014). The transposable element Bari-Jheh
78 mediates oxidative stress response in *Drosophila*. *Molecular Ecology* **23**, 2020–2030.
- 79 **Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K.** (2018). MEGA X: Molecular
80 Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35**, 1547–
81 1549.
- 82 **Lack, J. B., Lange, J. D., Tang, A. B., Corbett-Detig, R. B. and Pool, J. E.** (2016). A
83 Thousand Fly Genomes: An Expanded *Drosophila* Genome Nexus. *bioRxiv* 063537.
- 84 **Lavrinenko, A., Kesäniemi, J., Watts, P. C., Serga, S., Pascual, M., Mestres, F. and**
85 **Kozeretska, I.** (2017). First record of the invasive pest *Drosophila suzukii* in Ukraine
86 indicates multiple sources of invasion. *Journal of Pest Science* **90**, 421–429.
- 87 **Liguori, I., Russo, G., Curcio, F., Bulli, G., Aran, L., Della-Morte, D., Gargiulo, G., Testa, G.,**
88 **Cacciatore, F., Bonaduce, D., et al.** (2018). *Oxidative stress, aging, and diseases*. Dove
89 Medical Press Ltd.
- 90 **MacKay, T. F. C., Richards, S., Stone, E. A., Barbadilla, A., Ayroles, J. F., Zhu, D., Casillas,**
91 **S., Han, Y., Magwire, M. M., Cridland, J. M., et al.** (2012). The *Drosophila*
92 *melanogaster* Genetic Reference Panel. *Nature* **482**, 173–178vi.
- 93 **Marin, P., Genitoni, J., Barloy, D., Maury, S., Gibert, P., Ghalambor, C. K. and Vieira, C.**
94 (2019). Biological invasion: The influence of the hidden side of the (epi)genome.
95 *Functional Ecology* **34**, 385–400.
- 96 **Oesch, F., Herrero, M. E., Hengstler, J. G., Lohmann, M. and Arand, M.** (2000). Metabolic
97 Detoxification: Implications for Thresholds. *Toxicologic Pathology* **28**, 382–387.
- 98 **Parcy, F., Khan, A., Baranasic, D., Kulkarni, S. R., Stigliani, A., van der Lee, R.,**
99 **Vandepoele, K., Gheorghe, M., Lenhard, B., Tan, G., et al.** (2017). JASPAR 2018:
00 update of the open-access database of transcription factor binding profiles and its web
01 framework. *Nucleic Acids Research* **46**, D260–D266.
- 02 **Paris, M., Boyer, R., Jaenichen, R., Wolf, J., Karageorgi, M., Green, J., Cagnon, M.,**
03 **Parinello, H., Estoup, A., Gautier, M., et al.** (2020). Near-chromosome level genome
04 assembly of the fruit pest *Drosophila suzukii* using long-read sequencing. *bioRxiv*
05 2020.01.02.892844.
- 06 **R Core Team** (2019). R: A Language and Environment for Statistical Computing.
- 07 **Rzeczniczak, T. Z., Douglas, L. A., Watterson, J. H. and Merritt, T. J. S.** (2011). Paraquat
08 administration in *Drosophila* for use in metabolic studies of oxidative stress. *Analytical*
09 *Biochemistry* **419**, 345–347.
- 10 **Sandelin, A., Wasserman, W. W. and Lenhard, B.** (2004). ConSite: web-based prediction of
11 regulatory elements using cross-species comparison. *Nucleic Acids Res* **32**, W249–
12 W252.
- 13 **Sessegolo, C., Burlet, N. and Haudry, A.** (2016). Strong phylogenetic inertia on genome size
14 and transposable element content among 26 species of flies. *Biology Letters* **12**,
15 20160407.

- 16 **Shaw, M. W.** (2000). Models of the Effects of Dose Heterogeneity and Escape on Selection
17 Pressure for Pesticide Resistance. *Phytopathology*TM **90**, 333–339.
- 18 **Stapley, J., Santure, A. W. and Dennis, S. R.** (2015). Transposable elements as agents of
19 rapid adaptation may explain the genetic paradox of invasive species. *Molecular Ecology*
20 **24**, 2241–2252.
- 21 **Sylvain, L. P.** (2012). *EasyqpcR: EasyqpcR for low-throughput real-time quantitative PCR data*
22 *analysis*. Bioconductor version: Release (3.10).
- 23 **Tan, G. and Lenhard, B.** (2016). TFBSTools: an R/Bioconductor package for transcription
24 factor binding site analysis. *Bioinformatics* **32**, 1555–1556.
- 25 **Tsai, W.-T.** (2018). Status of herbicide use, regulatory management and case study of paraquat
26 in Taiwan. *Environment, Development and Sustainability*.
- 27 **Vermeulen, C. J., Van De Zande, L. and Bijlsma, R.** (2005). Resistance to Oxidative Stress
28 Induced by Paraquat Correlates Well with Both Decreased and Increased Lifespan in
29 *Drosophila melanogaster*. *Biogerontology* **6**, 387–395.
- 30 **Villanueva-Cañas, J. L., Horvath, V., Aguilera, L. and González, J.** (2019). Diverse families
31 of transposable elements affect the transcriptional regulation of stress-response genes in
32 *Drosophila melanogaster*. *Nucleic Acids Res* **47**, 6842–6857.
- 33 **Weber, A. L., Khan, G. F., Magwire, M. M., Tabor, C. L., Mackay, T. F. C. and Anholt, R. R.**
34 **H.** (2012). Genome-wide association analysis of oxidative stress resistance in *drosophila*
35 *melanogaster*. *PLoS ONE* **7**,.

36 List of the figures

37 Fig. 1. **Mean (\pm SD) of survival time (days) of *D. suzukii* genotypes (A) and reaction norm (log of the**
38 **mean) (B).** (A) mean values under control and paraquat condition for females (pink) and males (blue)
39 with associated SD. (B) reaction norm between treatments (control and paraquat) for both females and
40 males for the six genotypes using log transformed mean value. The slope differences between the
41 curves highlight the difference of sensitivity between genotypes.

42 Fig. 2. ***Jheh* gene expression (*Jheh-1, -2, -3*) after normalization by *Rp49* in control (green) and**
43 **paraquat (red) conditions for females on the left and males on the right.**

44 Fig. 3. **Representation of the *Jheh* gene cluster on the reference genome and the relative position on**
45 **the other six genotypes.** The blue squares represent the exons and the blue lines the intronic regions.
46 Red triangles represent TE detected in the intergenic regions and TFBS are indicated by vertical lines.
47 The scale is shown below the figure in base pairs.

48 Fig. 4. **TFBS detected in the four intergenic regions of the *Jheh* gene cluster in the reference genome**
49 **and in the six genotypes of *D. suzukii*.** Of the 14 TFBS examined, 8 had at least one positive result. The
50 intergenic regions are upstream *Jheh-1*(CG18190-*Jheh-1*), between *Jheh-1* and -2 (*Jheh1-Jheh2*),
51 between *Jheh2* and -3 (*Jheh2-Jheh3*) downstream of *Jheh-3* (*Jheh3-CG43069*).

52

53 List of the tables

54 Table 1. **Transcription factor analyzed with the PWM matrix ID from JASPAR2018.** Mainly matrix come
55 from *D. melanogaster* model, but several as HSF, HIF1 and XBP1 come from human, while MTF1 come
56 from mice.

57

58 Table 2. **Mean (\pm SD) survival time (days) for male and female lines of *D. suzukii* under**
59 **control and paraquat conditions.** Sensitivity represents the exponential of the interaction values
60 in the model (i.e., the difference in slope between the Sapporo reference and the other genotypes,
61 see Figure S1). * indicate a significant difference with the reference (p-value < 0.05).

62 Table 3. **Within-line diversity (π) for the six genotypes of *D. suzukii* for the seven sequenced intronic**
63 **regions.** The mean diversity per intron was calculated using the most common sequence of the six
64 genotypes.

65 Table 4. **Size differences (bp) between the six genotypes and the reference genome.** TE insertions are
66 indicated by their size or abs if they are absent.

67

68 List of supplementary figures

69 Fig. S1. **Representation of the parameters estimated by the model for treatment, genotypes and**
70 **interactions for females (left) and males (right).** Values were transformed exponentially to be
71 interpreted as a multiplicator effect. The vertical line corresponds to the reference. Associated p-values
72 are greater than 0.05 when the confidence interval includes the vertical line.

73 Fig. S2. **Distribution of the between genotype (blue) and within genotype (pink) genetic diversity using**
74 **π values for all intronic regions.**

75

76 List of supplementary tables

77

78 Table S1. **Primers used for the PCR experiments.**

79 Table S2. **Genotypes of *D. suzukii* with the number of haplotypes obtained.** The length of the sequence
80 corresponds to the size of the amplified fragment used to calculate the average diversity per population
81 per intron. The mean diversity was calculated as the average between the most common haplotypes of
82 the 6 lines.

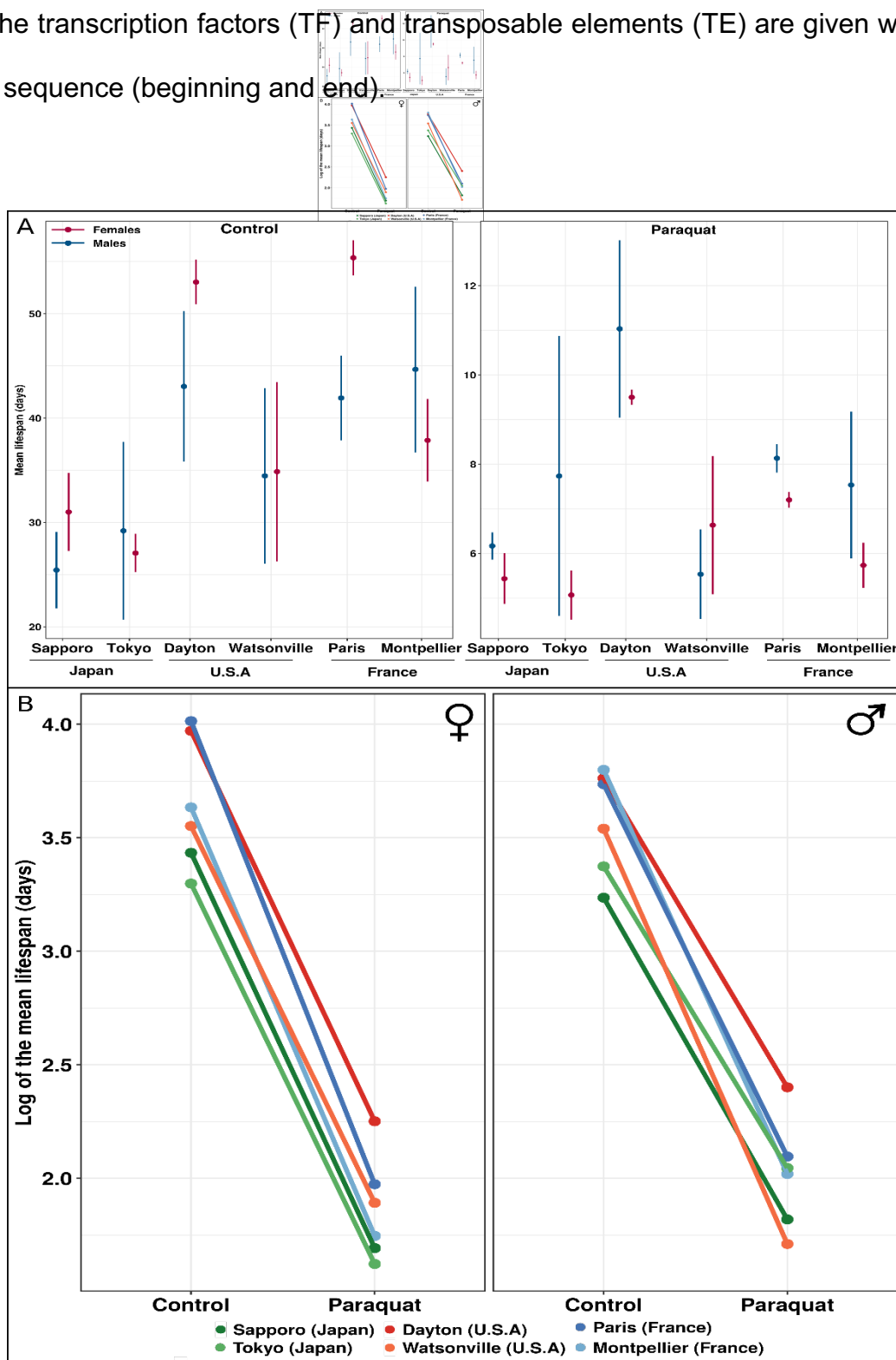
83 Table S3. **Pairwise genetic distance between genotypes. Bold values represent genetic diversity within**
84 **lines.**

85 Table S4. **Summary of the ANOVA 2 by gene and sex.**

86

87 **Table S5. TFBS and TE detected in all six genotypes and the reference genome of *D.***
 88 ***suzukii*.** We screened TFBS in the intergenic regions before, between and after the Jheh genes.
 89 The names of the transcription factors (TF) and transposable elements (TE) are given with their
 90 positions in the sequence (beginning and end).

91



92

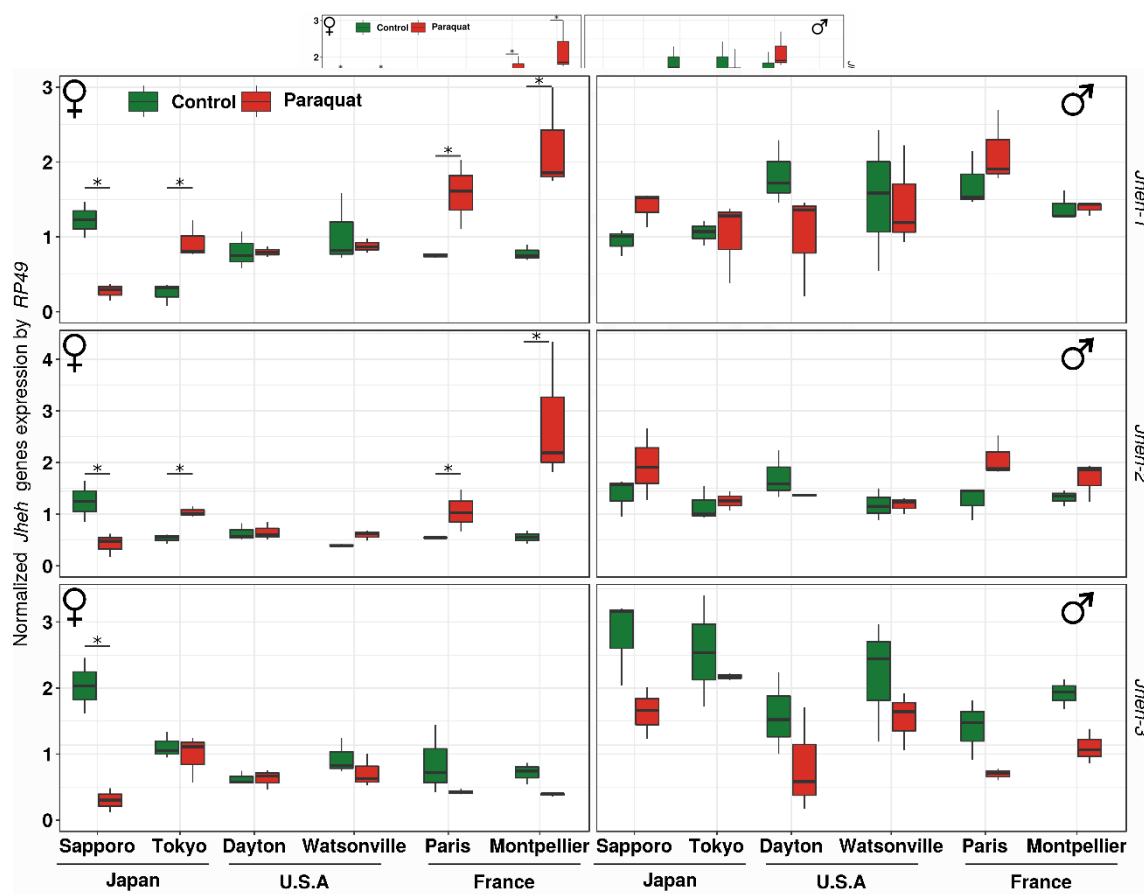
93

94

95 Figure 1

96

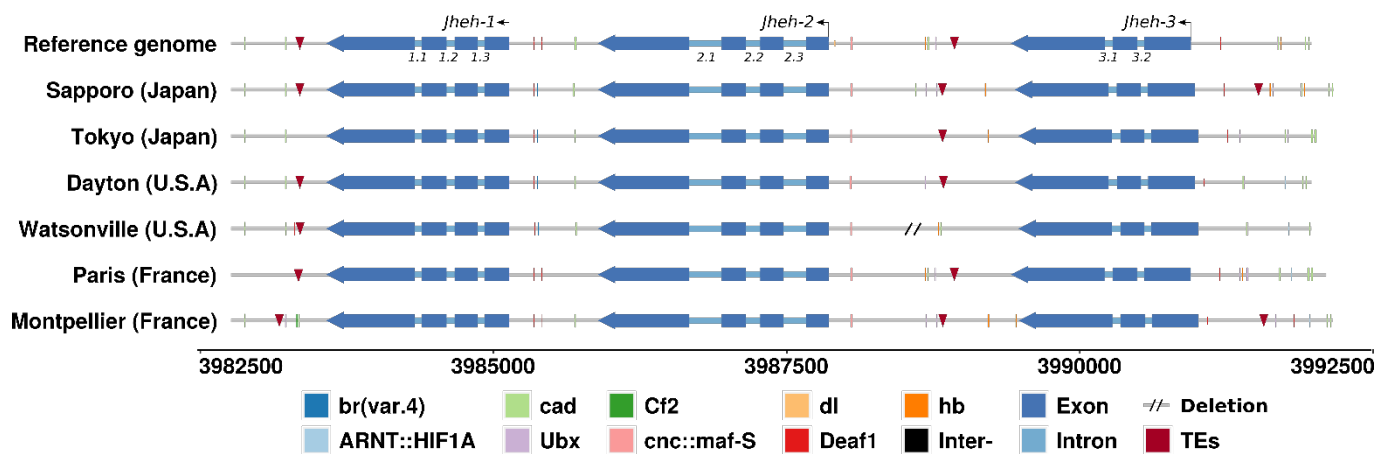
97
98
99
00
01



02
03
04
05
06
07
08
09
10

Figure 2

11



12

13 Figure 3

14

15

16

17

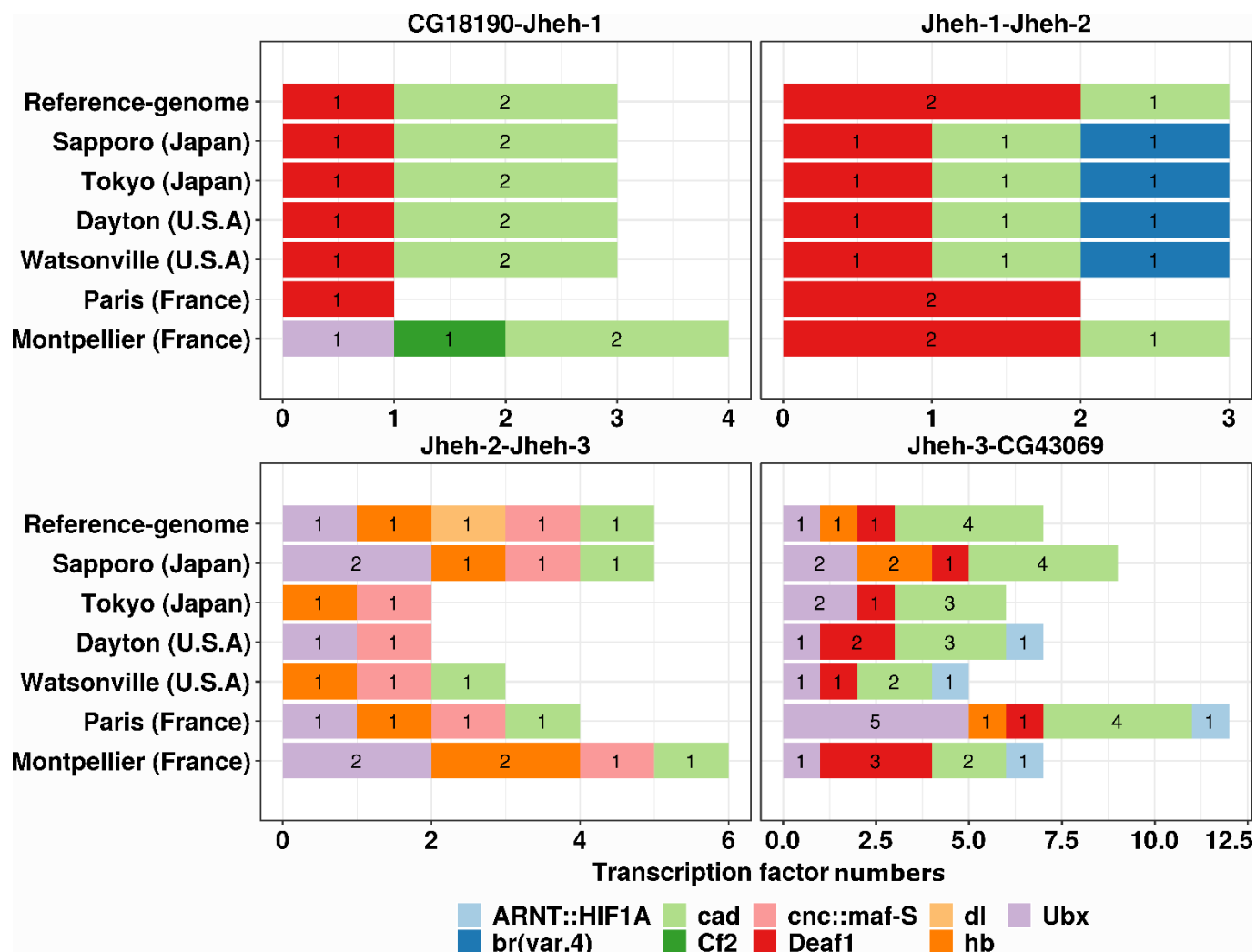
18

19

20

21

22



23 Figure 4

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38 Table 1. **Transcription factor analyzed with the PWM matrix ID from JASPAR2018.** Mainly
 39 matrix come from *D. melanogaster* model, but several as HSF, HIF1 and XBP1 come from human,
 40 while MTF1 come from mice.

Trancription factors	PWM ID	Species	Origin
HSF (heat shock factor)	MA0486.2	<i>Homo sapiens</i>	
HIF1 (hypoxia inducible factor)	MA0259.1	<i>Homo sapiens</i>	
DL (Dorsal)	MA0022.1	<i>D. melanogaster</i>	
MTF1 (Metal response element-binding Transcription Factor-1)	PB0044.1	<i>Mus musculus</i>	
DEAF1 (Deformed epidermal autoregulatory factor-1)	MA0185.1	<i>D. melanogaster</i>	From Villanueva-Cañas <i>et al.</i> (2019)
CAD (caudal)	MA0216.2	<i>D. melanogaster</i>	
NUB (nubbin)	MA0197.2	<i>D. melanogaster</i>	
XBP1 (X box binding protein-1)	MA0844.1	<i>Homo sapiens</i>	
CnC (cap-n-collar)	MA0530.1	<i>D. melanogaster</i>	
Br(var4) (broad complex 4)	MA0013.1	<i>D. melanogaster</i>	
Hb (hunchbak)	MA0049.1	<i>D. melanogaster</i>	
Ubx (Ultrabithorax)	MA0094.2	<i>D. melanogaster</i>	From conSite website (Sandelin <i>et al.</i> , 2004)
Cf2 (Chorion factor 2)	MA0015.1	<i>D. melanogaster</i>	
Snail (sna)	MA0086.2	<i>D. melanogaster</i>	

41
 42
 43

44 Table 2. **Mean (\pm SD) survival time (days) for male and female lines of *D. suzukii* under**
 45 **control and paraquat conditions.** Sensitivity represents the exponential of the interaction values
 46 in the model (i.e., the difference in slope between the Sapporo reference and the other genotypes,
 47 see Figure S1). * indicate a significant difference with the reference (p-value < 0.05).

Lines	Females			Males		
	Control	Paraquat	sensitivity	Control	Paraquat	sensitivity
Sapporo (Japan)	31 \pm 13.3	5.4 \pm 2.3	1.0	25.4 \pm 5.9	6.2 \pm 2.1	1.0
Tokyo (Japan)	27.1 \pm 9.1	5.1 \pm 1.9	1.06	29.2 \pm 11.5	7.7 \pm 4.9	1.02
Dayton (U.S.A)	53 \pm 8.4	9.5 \pm 3	0.98	42.8 \pm 16.3	11 \pm 4.1	1.23
Watsonville (U.S.A)	34.9 \pm 13.8	6.6 \pm 2.3	1.13	34.5 \pm 11.2	5.5 \pm 2.1	0.68*
Paris (France)	55.4 \pm 9.4	7.2 \pm 1.8	0.72*	41.9 \pm 14.8	8.1 \pm 2.8	0.9
Montpellier (France)	37.9 \pm 8.4	5.7 \pm 2.2	0.83	44.7 \pm 12.2	7.5 \pm 3.8	0.66*

48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69

70 **Table 3 Within-line diversity (pi) for the six genotypes of *D. sukuzii* for the seven sequenced intronic**
71 **regions.** The mean diversity per intron was calculated using the most common sequence of the six
72 genotypes.

	<i>Jheh-1.1</i>	<i>Jheh-1.2</i>	<i>Jheh-1.3</i>	<i>Jheh-2.1</i>	<i>Jheh-2.2</i>	<i>Jheh-2.3</i>	<i>Jheh-3.2</i>
Paris	0.0000	0.0000	0.0000	0.0460	0.0000	0.0000	0.0000
Montpellier	0.0000	0.0169	0.0000	0.0502	0.0081	0.0000	0.0339
Sapporo	0.0000	0.0000	0.0396	0.0546	0.0000	0.0000	0.0000
Tokyo	0.0000	0.0000	0.0000	0.0324	0.0000	0.0144	0.0113
Dayton	0.0000	0.0000	0.0000	0.0449	0.0000	0.0096	0.0000
Watsonville	0.0792	0.0000	0.0198	0.0466	0.0000	0.0000	0.0000
Mean pi diversity	0.0070	0.0080	0.0177	0.0390	0.0214	0.0333	0.0209

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

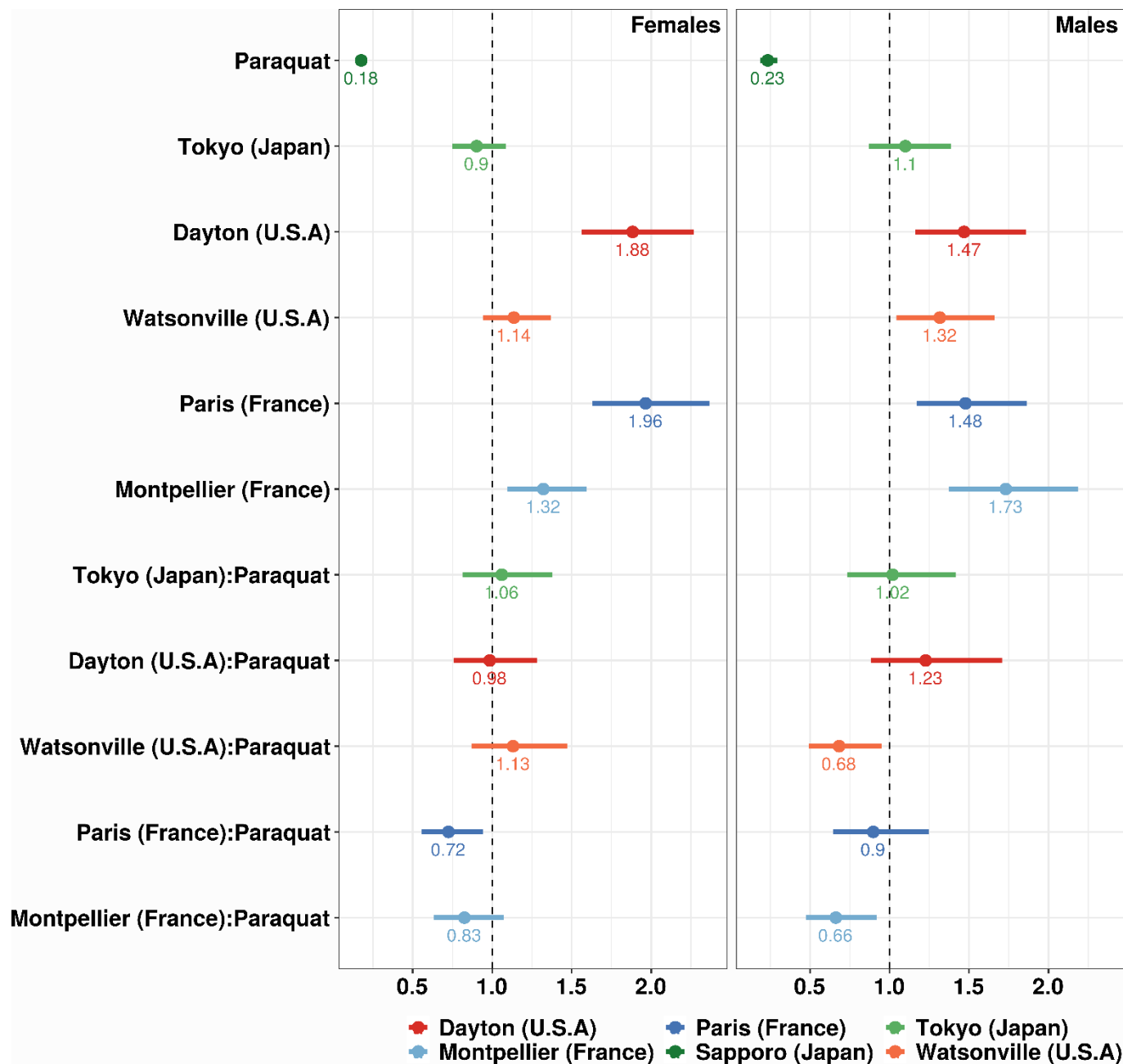
96

97

98 **Table 4. Size differences (bp) between the six genotypes and the reference genome. TE**
 99 insertions are indicated by their size or abs if they are absent.

		expected size (reference genome)	Sapporo (Japan)	Tokyo (Japan)	Dayton (U.S.A)	Watsonville (U.S.A)	Paris (France)	Montpellier (France)
X-Jheh1	Size	809	0	-2	0	+1	0	+12
	TE insertion	82	82	abs	82	82	82	abs
Jheh1/jheh2	Size	756	-17	-6	-2	+2	+1	+7
	TE insertion	abs	abs	abs	abs	abs	abs	abs
Jheh2/Jheh3	Size	1541	+31	-64	+35	-297	-1	+66
	TE insertion	41	48	49	91	abs	41	49
Jheh3-Y	Size	1021	+171	-8	+58	-55	+29	+122
	TE insertion	abs	58	abs	abs	abs	abs	36

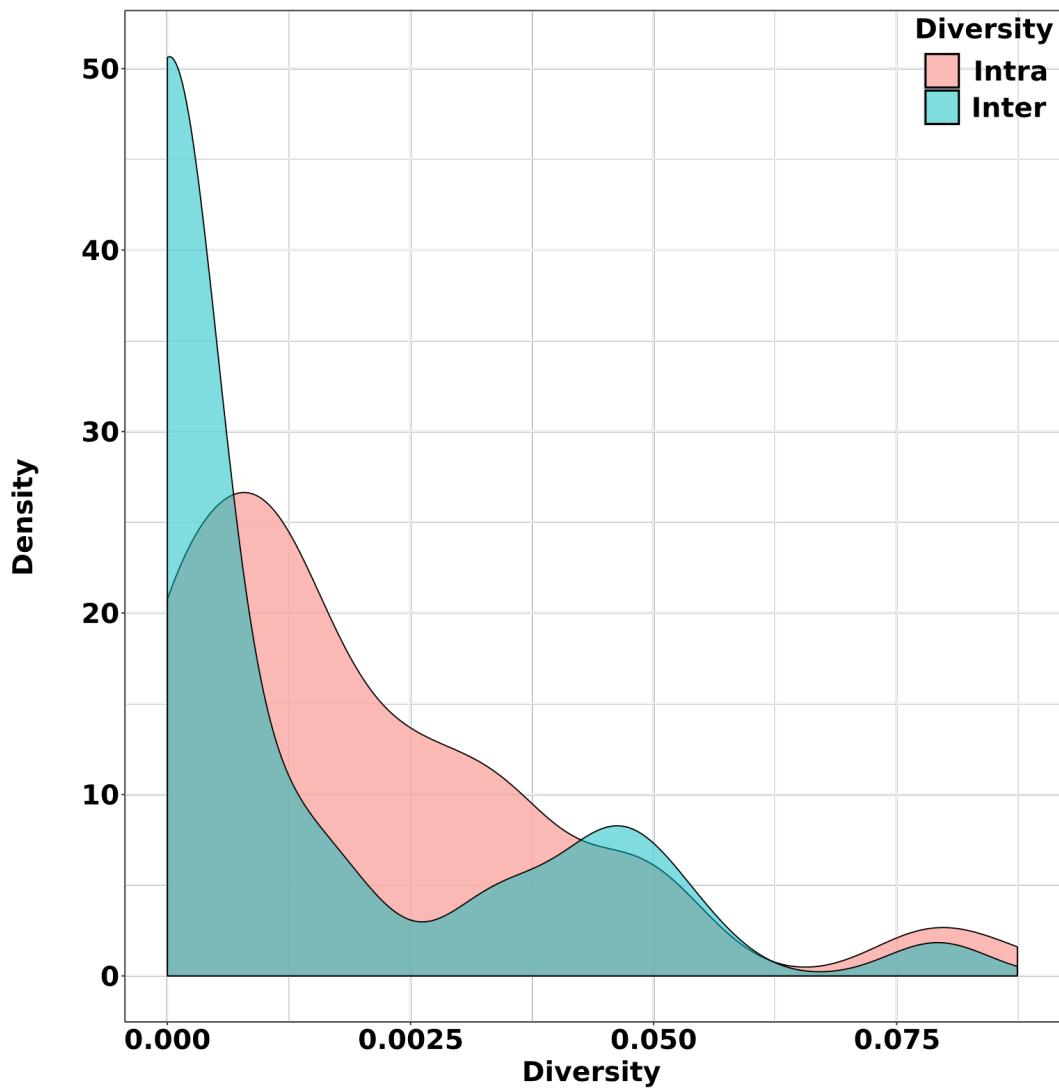
00
 01
 02
 03
 04
 05
 06
 07



08
09
10

11 Fig. S1. Representation of the parameters estimated by the model for treatment, genotypes and
12 interactions for females (left) and males (right). Values were transformed exponentially to be
13 interpreted as a multiplier effect. The vertical line corresponds to the reference. Associated p-values
14 are greater than 0.05 when the confidence interval includes the vertical line.

15
16



17

18 Fig. S2. Distribution of the between genotype (blue) and within genotype (pink) genetic diversity using
19 pi values for all intronic regions.

20

21

22

23

24

25

26

27

28

29

30 Table S1. Primers used for the PCR experiments.

31

Name	Sequence
<i>Jheh-1</i>	Forward: GAGCAACCTGGACAAGAACAAC Reverse: TATCCCAAGCGCTGCATAAG
<i>Jheh-2</i>	Forward: AGAAGCTGGACCACTACCAAAC Reverse: AGAACCTTCTTGGGCTTCTGG
<i>Jheh-3</i>	Forward: AGTACGCTTTTGAGGTCGTG Reverse: AGACGCAGCATCAAGTTTCG
<i>RP49</i>	Forward: CCGCTTCAAGGGACAGTATC Reverse: GACGATCTCCTTGCCTTCT
<i>DS10_00005800-Jheh-1</i>	Forward: GTGTCCCTGGACCATGTTGT Reverse: GGAGGACACTTTGCGGCTAT
<i>Jheh-1-Jheh-2</i>	Forward: GCCAATGGCCAGTACACAGA Reverse: GCCCCAGAAGCTGTACGATG
<i>Jheh-2-Jheh-3</i>	Forward: GCAAAGTGAGCATGATTTGGC Reverse: CAACCCTGTGAACCGAGCTA
<i>DS10_0005804-Jheh-3</i>	Forward: GCAATTAGCTCCCACTCGGT Reverse: CGTGACACTGCAGTTTATGGC
<i>Jheh-1_intr1</i>	Forward: GAGCGGATCCTAGACCCTTC Reverse: GCTGGTCGGAGGTAAGTTGT
<i>Jheh-1_intr2</i>	Forward: AAGAAAGTGCATGCGTAGCC Reverse: TGGCAGTTCAACCACTTCAC
<i>Jheh-1_intr3</i>	Forward: ATTGAGGCGGCTCTTTAGGT Reverse: CGGAGGTGATAAACAACAACTT
<i>Jheh-2_intr1</i>	Forward: GAGGCCTGGAATTGGAAAAT Reverse: TCTCGAGGGAATAAGAGGTTCA
<i>Jheh-2_intr2</i>	Forward: CGGCTTGGCATGAATAAAGT Reverse: ACGGAGATCCAGGGGTAAGT
<i>Jheh-2_intr3</i>	Forward: CCTCAATTACCTGTGGGGTAAA Reverse: CCCGAGGTAAGCTATGTTTCA
<i>Jheh-3_intr2</i>	Forward: GCCTTCTCGTGAACGTAGTGA Reverse: CAAGCAGTACACGACCGAGA

32

33

34

35

36

37

38 Table S2. **Genotypes of *D. suzukii* with the number of haplotypes obtained.** The length of
 39 the sequence corresponds to the size of the amplified fragment used to calculate the average
 40 diversity per population per intron. The mean diversity was calculated as the average between
 41 the most common haplotypes of the 6 lines.
 42

Lineage	Jheh-1 intron 1			
	# of haplotypes	# of flies	Size of the alignment	Mean diversity
Paris (France)	1	6	97pb	0.0070
Montpellier (France)	1	10		
Sapporo Japan)	1	6		
Tokyo (Japan)	1	10		
Dayton (U.S.A)	1	10		
Watsonville (U.S.A)	2	7-2		
	Jheh-1 intron 2			
	N haplotype	N	Alignment	Mean diversity
Paris (France)	1	10	119pb	0.0080
Montpellier (France)	2	4-3		
Sapporo Japan)	1	10		
Tokyo (Japan)	1	10		
Dayton (U.S.A)	1	10		
Watsonville (U.S.A)	1	10		
	Jheh-1 intron 3			
	N haplotype	N	Alignment	Mean diversity

Paris (France)	1	10	133pb	0.0177
Montpellier (France)	1	10		
Sapporo Japan)	2	7-2		
Tokyo (Japan)	1	10		
Dayton (U.S.A)	1	10		
Watsonville (U.S.A)	2	5-5		
		Jheh-2 intron 1		
	N haplotype	N	Alignment	Mean diversity
Paris (France)	2	4-2	250pb	0.0390
Montpellier (France)	2	5-2		
Sapporo Japan)	2	4-3		
Tokyo (Japan)	4	2-2-1-1		
Dayton (U.S.A)	3	4-2-1		
Watsonville (U.S.A)	2	5-4		
		Jheh-2 intron 2		
	N haplotype	N	Alignment	Mean diversity
Paris (France)	1	8	165pb	0.0214
Montpellier (France)	3	4-3-2		
Sapporo Japan)	1	10		
Tokyo (Japan)	1	8		

Dayton (U.S.A)	1	9		
Watsonville (U.S.A)	1	10		
		Jheh-2 intron 3		
	N haplotype	N	Alignment	Mean diversity
Paris (France)	1	8	211pb	0.0333
Montpellier (France)	1	8		
Sapporo Japan)	1	10		
Tokyo (Japan)	2	5-4		
Dayton (U.S.A)	2	4-2		
Watsonville (U.S.A)	1	10		
		Jheh-3 intron 2		
	N haplotype	N	Alignment	Mean diversity
Paris (France)	1	10	118pb	0.0209
Montpellier (France)	2	7-3		
Sapporo Japan)	1	9		
Tokyo (Japan)	3	7-2-1		
Dayton (U.S.A)	1	9		
Watsonville (U.S.A)	1	10		

43

44 Table S3. Pairwise genetic distance between genotypes. Bold values represent genetic diversity within
45 lines.

46

Jheh-1.1	Paris	Montpellier	Sapporo	Tokyo	Dayton	Watsonville
Paris	0.0000					
Montpellier	0.0104	0.0000				
Sapporo	0.0104	0.0000	0.0000			
Tokyo	0.0208	0.0104	0.0104	0.0000		
Dayton	0.0105	0.0000	0.0000	0.0105	0.0000	
Watsonville	0.0104	0.0000	0.0000	0.0104	0.0000	0.0792
Jheh-1.2	Paris	Montpellier	Sapporo	Tokyo	Dayton	Watsonville
Paris	0.0000					
Montpellier	0.0084	0.0169				
Sapporo	0.0000	0.0084	0.0000			
Tokyo	0.0000	0.0084	0.0000	0.0000		
Dayton	0.0000	0.0084	0.0000	0.0000	0.0000	
Watsonville	0.0000	0.0084	0.0000	0.0000	0.0000	0.0000
Jheh-1.3	Paris	Montpellier	Sapporo	Tokyo	Dayton	Watsonville
Paris	0.0000					
Montpellier	0.0079	0.0000				
Sapporo	0.0157	0.0236	0.0396			
Tokyo	0.0079	0.0000	0.0236	0.0000		
Dayton	0.0238	0.0159	0.0379	0.0159	0.0000	
Watsonville	0.0236	0.0157	0.0226	0.0157	0.0152	0.0198
Jheh-2.1	Paris	Montpellier	Sapporo	Tokyo	Dayton	Watsonville
Paris	0.0460					
Montpellier	0.0502	0.0502				
Sapporo	0.0546	0.0462	0.0546			
Tokyo	0.0502	0.0167	0.0378	0.0324		

Dayton	0.0254	0.0466	0.0511	0.0466	0.0449	
Watsonville	0.0502	0.0084	0.0462	0.0084	0.0466	0.0466
Jheh-2.2	Paris	Montpellier	Sapporo	Tokyo	Dayton	Watsonville
Paris	0.0000					
Montpellier	0.0121	0.0081				
Sapporo	0.0303	0.0303	0.0000			
Tokyo	0.0061	0.0061	0.0242	0.0000		
Dayton	0.0364	0.0364	0.0242	0.0303	0.0000	
Watsonville	0.0000	0.0121	0.0303	0.0061	0.0364	0.0000
Jheh-2.3	Paris	Montpellier	Sapporo	Tokyo	Dayton	Watsonville
Paris	0.0000					
Montpellier	0.0240	0.0000				
Sapporo	0.0144	0.0096	0.0000			
Tokyo	0.0144	0.0096	0.0000	0.0144		
Dayton	0.0144	0.0096	0.0000	0.0000	0.0096	
Watsonville	0.0825	0.0874	0.0777	0.0777	0.0777	0.0000
Jheh-3.2	Paris	Montpellier	Sapporo	Tokyo	Dayton	Watsonville
Paris	0.0000					
Montpellier	0.0339	0.0339				
Sapporo	0.0254	0.0085	0.0000			
Tokyo	0.0339	0.0169	0.0085	0.0113		
Dayton	0.0000	0.0339	0.0254	0.0339	0.0000	
Watsonville	0.0339	0.0169	0.0085	0.0000	0.0339	0.0000

47

48 **Table S4. Summary of the ANOVA 2 by gene and sex.**

49

50

51

Gene	Sex	Estimate	Df	Sum Sq	Mean Sq	F value	Pr(>F)
<i>Jheh-1</i>	Females	Treatment	1	0.7578741	0.7578741	5.4488993	0.0286769
		Genotype	5	5.7726966	1.1545393	8.3008099	0.0001334
		Treatment:Genotype	5	7.6349146	1.5269829	10.9785737	0.0000172
		Residuals	23	3.1990137	0.1390876	NA	NA
	Males	Treatment	1	0.0375774	0.0375774	0.1612375	0.6915729
		Genotype	5	1.5806168	0.3161234	1.3564255	0.2754043
		Treatment:Genotype	5	1.4780750	0.2956150	1.2684280	0.3096968
		Residuals	24	5.5933488	0.2330562	NA	NA
<i>Jheh-2</i>	Females	Treatment	1	1.3696002	1.3696002	13.3644443	0.0013171
		Genotype	5	2.7850846	0.5570169	5.4353244	0.0019155
		Treatment:Genotype	5	5.2867104	1.0573421	10.3174554	0.0000276
		Residuals	23	2.3570606	0.1024809	NA	NA
	Males	Treatment	1	0.2389486	0.2389486	4.2568387	0.0500705
		Genotype	5	0.6101519	0.1220304	2.1739557	0.0908517
		Treatment:Genotype	5	0.4690600	0.0938120	1.6712490	0.1799009
		Residuals	24	1.3471889	0.0561329	NA	NA
<i>Jheh-3</i>	Females	Treatment	1	2.8083330	2.8083330	21.6508485	0.0001105
		Genotype	5	1.8020725	0.3604145	2.7786163	0.0417902
		Treatment:Genotype	5	3.4117019	0.6823404	5.2605045	0.0022994
		Residuals	23	2.9833315	0.1297101	NA	NA
	Males	Treatment	1	2.5825122	2.5825122	13.5076975	0.0011917
		Genotype	5	4.5987363	0.9197473	4.8106909	0.0034625
		Treatment:Genotype	5	0.6865256	0.1373051	0.7181674	0.6161005
		Residuals	24	4.5885164	0.1911882	NA	NA

52
53
54

55 Table S5. **TFBS and TE detected in all six genotypes and the reference genome of *D.***
56 ***suzukii***. We screened TFBS in the intergenic regions before, between and after the Jheh genes.
57 The names of the transcription factors (TF) and transposable elements (TE) are given with their
58 positions in the sequence (beginning and end).

Genotype	Gene	subgene	Element	Start	end
Dayton (U.S.A)	CG18190-Jheh-1	RLX- incomp_Blc1935_Dsuz- L-B2033- Map1_reversed	TE	3983310	3983388
Montpellier (France)	CG18190-Jheh-1	RXX_Blc1636_Dsuz-B- R3220- Map20_reversed	TE	3983164	3983184
Paris (France)	CG18190-Jheh-1	RLX- incomp_Blc1935_Dsuz- L-B2033- Map1_reversed	TE	3983310	3983388
Reference genome	CG18190-Jheh-1	RLX- incomp_Blc1935_Dsuz- L-B2033- Map1_reversed	TE	3983310	3983388
Sapporo (Japan)	CG18190-Jheh-1	RLX- incomp_Blc1935_Dsuz- L-B2033- Map1_reversed	TE	3983310	3983388
Tokyo (Japan)	CG18190-Jheh-1	RXX- LARD_Blc2842_Dsuz- L-B3109- Map1_reversed	TE	3983339	3983380
Watsonville (U.S.A)	CG18190-Jheh-1	RLX- incomp_Blc1935_Dsuz- L-B2033- Map1_reversed	TE	3983312	3983390
Dayton (U.S.A)	Jheh-2-Jheh-3	RXX- LARD_Blc2479_Dsuz- L-B2652-Map1	TE	3988791	3988881
Montpellier (France)	Jheh-2-Jheh-3	RXX- LARD_Blc2479_Dsuz- L-B2652-Map1	TE	3988807	3988855
Paris (France)	Jheh-2-Jheh-3	RLX-incomp- chim_Blc427_Dsuz-L- B425-Map1	TE	3988909	3988949
Reference genome	Jheh-2-Jheh-3	RLX-incomp- chim_Blc427_Dsuz-L- B425-Map1	TE	3988910	3988950
Sapporo (Japan)	Jheh-2-Jheh-3	RXX- LARD_Blc2479_Dsuz- L-B2652-Map1	TE	3988804	3988850
Tokyo (Japan)	Jheh-2-Jheh-3	RXX- LARD_Blc2479_Dsuz- L-B2652-Map1	TE	3988805	3988853
Dayton (U.S.A)	Jheh-3-CG43069	RXX- LARD_Blc5020_Dsuz- L-B5102-Map1	TE	3991489	3991526
Montpellier (France)	Jheh-3-CG43069	RXX- LARD_Blc4946_Dsuz- L-B5036-Map1	TE	3991470	3991507
Reference genome	Jheh-3-CG43069	RXX-LARD- chim_Blc2440_Dsuz-L-	TE	3991670	3991701

B2608-Map1					
Sapporo (Japan)	Jheh-3-CG43069	RXX- incomp_Blc4021_Dsuz- L-B4274-Map1	TE	3991449	3991507
Tokyo (Japan)	Jheh-3-CG43069	DHX- incomp_Blc652_Dsuz- B-R2897-Map20	TE	3991666	3991691
Watsonville (U.S.A)	Jheh-3-CG43069	RXX_Blc2359_Dsuz-B- R8146-Map6_reversed	TE	3991575	3991608
Paris (France)	Jheh-3-CG43069	RXX- LARD_Blc434_Dsuz-L- B429-Map1	TE	3991559	3991592
Dayton (U.S.A)	CG18190-Jheh-1	cad	TF	3983223	3983233
Dayton (U.S.A)	CG18190-Jheh-1	cad	TF	3982874	3982884
Dayton (U.S.A)	CG18190-Jheh-1	Deaf1	TF	3983298	3983303
Montpellier (France)	CG18190-Jheh-1	cad	TF	3983335	3983345
Montpellier (France)	CG18190-Jheh-1	cad	TF	3982874	3982884
Montpellier (France)	CG18190-Jheh-1	Ubx	TF	3983223	3983230
Montpellier (France)	CG18190-Jheh-1	Cf2	TF	3983319	3983328
Paris (France)	CG18190-Jheh-1	Deaf1	TF	3983298	3983303
Reference genome	CG18190-Jheh-1	cad	TF	3983223	3983233
Reference genome	CG18190-Jheh-1	cad	TF	3982874	3982884
Reference genome	CG18190-Jheh-1	Deaf1	TF	3983298	3983303
Sapporo (Japan)	CG18190-Jheh-1	cad	TF	3983223	3983233
Sapporo (Japan)	CG18190-Jheh-1	cad	TF	3982874	3982884
Sapporo (Japan)	CG18190-Jheh-1	Deaf1	TF	3983298	3983303
Tokyo (Japan)	CG18190-Jheh-1	cad	TF	3983223	3983233
Tokyo (Japan)	CG18190-Jheh-1	cad	TF	3982874	3982884
Tokyo (Japan)	CG18190-Jheh-1	Deaf1	TF	3983298	3983303
Watsonville (U.S.A)	CG18190-Jheh-1	cad	TF	3983225	3983235
Watsonville (U.S.A)	CG18190-Jheh-1	cad	TF	3982874	3982884
Watsonville (U.S.A)	CG18190-Jheh-1	Deaf1	TF	3983300	3983305
Dayton	Jheh-1-Jheh-2	cad	TF	3985692	3985702

(U.S.A)						
Dayton (U.S.A)	Jheh-1-Jheh-2	Deaf1	TF	3985342	3985347	
Dayton (U.S.A)	Jheh-1-Jheh-2	br(var.4)	TF	3985371	3985381	
Montpellier (France)	Jheh-1-Jheh-2	cad	TF	3985692	3985702	
Montpellier (France)	Jheh-1-Jheh-2	Deaf1	TF	3985341	3985346	
Montpellier (France)	Jheh-1-Jheh-2	Deaf1	TF	3985407	3985412	
Paris (France)	Jheh-1-Jheh-2	Deaf1	TF	3985342	3985347	
Paris (France)	Jheh-1-Jheh-2	Deaf1	TF	3985408	3985413	
Reference genome	Jheh-1-Jheh-2	cad	TF	3985694	3985704	
Reference genome	Jheh-1-Jheh-2	Deaf1	TF	3985343	3985348	
Reference genome	Jheh-1-Jheh-2	Deaf1	TF	3985409	3985414	
Sapporo (Japan)	Jheh-1-Jheh-2	cad	TF	3985680	3985690	
Sapporo (Japan)	Jheh-1-Jheh-2	Deaf1	TF	3985341	3985346	
Sapporo (Japan)	Jheh-1-Jheh-2	br(var.4)	TF	3985370	3985380	
Tokyo (Japan)	Jheh-1-Jheh-2	cad	TF	3985691	3985701	
Tokyo (Japan)	Jheh-1-Jheh-2	Deaf1	TF	3985341	3985346	
Tokyo (Japan)	Jheh-1-Jheh-2	br(var.4)	TF	3985370	3985380	
Watsonville (U.S.A)	Jheh-1-Jheh-2	cad	TF	3985700	3985710	
Watsonville (U.S.A)	Jheh-1-Jheh-2	Deaf1	TF	3985350	3985355	
Watsonville (U.S.A)	Jheh-1-Jheh-2	br(var.4)	TF	3985379	3985389	
Dayton (U.S.A)	Jheh-2-Jheh-3	cnc::maf-S	TF	3988041	3988055	
Dayton (U.S.A)	Jheh-2-Jheh-3	Ubx	TF	3988678	3988685	
Montpellier (France)	Jheh-2-Jheh-3	cad	TF	3989450	3989460	
Montpellier (France)	Jheh-2-Jheh-3	cnc::maf-S	TF	3988043	3988057	
Montpellier (France)	Jheh-2-Jheh-3	hb	TF	3989217	3989226	
Montpellier (France)	Jheh-2-Jheh-3	hb	TF	3989452	3989461	
Montpellier (France)	Jheh-2-Jheh-3	Ubx	TF	3988684	3988691	

Montpellier (France)	Jheh-2-Jheh-3	Ubx	TF	3988777	3988784
Paris (France)	Jheh-2-Jheh-3	cad	TF	3988701	3988711
Paris (France)	Jheh-2-Jheh-3	cnc::maf-S	TF	3988043	3988057
Paris (France)	Jheh-2-Jheh-3	hb	TF	3988679	3988688
Paris (France)	Jheh-2-Jheh-3	Ubx	TF	3988764	3988771
Reference genome	Jheh-2-Jheh-3	dl	TF	3987904	3987915
Reference genome	Jheh-2-Jheh-3	cad	TF	3988702	3988712
Reference genome	Jheh-2-Jheh-3	cnc::maf-S	TF	3988044	3988058
Reference genome	Jheh-2-Jheh-3	hb	TF	3988680	3988689
Reference genome	Jheh-2-Jheh-3	Ubx	TF	3988765	3988772
Sapporo (Japan)	Jheh-2-Jheh-3	cad	TF	3988594	3988604
Sapporo (Japan)	Jheh-2-Jheh-3	cnc::maf-S	TF	3988043	3988057
Sapporo (Japan)	Jheh-2-Jheh-3	hb	TF	3989189	3989198
Sapporo (Japan)	Jheh-2-Jheh-3	Ubx	TF	3988683	3988690
Sapporo (Japan)	Jheh-2-Jheh-3	Ubx	TF	3988774	3988781
Tokyo (Japan)	Jheh-2-Jheh-3	cnc::maf-S	TF	3988042	3988056
Tokyo (Japan)	Jheh-2-Jheh-3	hb	TF	3989215	3989224
Watsonville (U.S.A)	Jheh-2-Jheh-3	cad	TF	3988809	3988819
Watsonville (U.S.A)	Jheh-2-Jheh-3	cnc::maf-S	TF	3988042	3988056
Watsonville (U.S.A)	Jheh-2-Jheh-3	hb	TF	3988790	3988799
Dayton (U.S.A)	Jheh-3-CG43069	ARNT::HIF1A	TF	3991698	3991705
Dayton (U.S.A)	Jheh-3-CG43069	cad	TF	3991337	3991347
Dayton (U.S.A)	Jheh-3-CG43069	cad	TF	3991874	3991884
Dayton (U.S.A)	Jheh-3-CG43069	cad	TF	3991845	3991855
Dayton (U.S.A)	Jheh-3-CG43069	Deaf1	TF	3991182	3991187
Dayton (U.S.A)	Jheh-3-CG43069	Deaf1	TF	3991006	3991011
Dayton	Jheh-3-CG43069	Ubx	TF	3991351	3991358

(U.S.A)						
Montpellier (France)	Jheh-3-CG43069	ARNT::HIF1A	TF	3991877	3991884	
Montpellier (France)	Jheh-3-CG43069	cad	TF	3992054	3992064	
Montpellier (France)	Jheh-3-CG43069	cad	TF	3992024	3992034	
Montpellier (France)	Jheh-3-CG43069	Deaf1	TF	3991182	3991187	
Montpellier (France)	Jheh-3-CG43069	Deaf1	TF	3991744	3991749	
Montpellier (France)	Jheh-3-CG43069	Deaf1	TF	3991006	3991011	
Montpellier (France)	Jheh-3-CG43069	Ubx	TF	3991589	3991596	
Reference genome	Jheh-3-CG43069	cad	TF	3991933	3991943	
Reference genome	Jheh-3-CG43069	cad	TF	3991674	3991684	
Reference genome	Jheh-3-CG43069	cad	TF	3991696	3991706	
Reference genome	Jheh-3-CG43069	cad	TF	3991903	3991913	
Reference genome	Jheh-3-CG43069	Deaf1	TF	3991182	3991187	
Reference genome	Jheh-3-CG43069	hb	TF	3991695	3991704	
Reference genome	Jheh-3-CG43069	Ubx	TF	3991667	3991674	
Sapporo (Japan)	Jheh-3-CG43069	cad	TF	3992103	3992113	
Sapporo (Japan)	Jheh-3-CG43069	cad	TF	3991844	3991854	
Sapporo (Japan)	Jheh-3-CG43069	cad	TF	3991866	3991876	
Sapporo (Japan)	Jheh-3-CG43069	cad	TF	3992073	3992083	
Sapporo (Japan)	Jheh-3-CG43069	Deaf1	TF	3991182	3991187	
Sapporo (Japan)	Jheh-3-CG43069	hb	TF	3991572	3991581	
Sapporo (Japan)	Jheh-3-CG43069	hb	TF	3991865	3991874	
Sapporo (Japan)	Jheh-3-CG43069	Ubx	TF	3991597	3991604	
Sapporo (Japan)	Jheh-3-CG43069	Ubx	TF	3991837	3991844	
Tokyo (Japan)	Jheh-3-CG43069	cad	TF	3991924	3991934	
Tokyo (Japan)	Jheh-3-CG43069	cad	TF	3991665	3991675	
Tokyo (Japan)	Jheh-3-CG43069	cad	TF	3991894	3991904	

Tokyo (Japan)	Jheh-3-CG43069	Deaf1	TF	3991176	3991181
Tokyo (Japan)	Jheh-3-CG43069	Ubx	TF	3991688	3991695
Tokyo (Japan)	Jheh-3-CG43069	Ubx	TF	3991281	3991288
Watsonville (U.S.A)	Jheh-3-CG43069	ARNT::HIF1A	TF	3991698	3991705
Watsonville (U.S.A)	Jheh-3-CG43069	cad	TF	3991337	3991347
Watsonville (U.S.A)	Jheh-3-CG43069	cad	TF	3991877	3991887
Watsonville (U.S.A)	Jheh-3-CG43069	Deaf1	TF	3991182	3991187
Watsonville (U.S.A)	Jheh-3-CG43069	Ubx	TF	3991351	3991358
Paris (France)	Jheh-3-CG43069	ARNT::HIF1A	TF	3991784	3991791
Paris (France)	Jheh-3-CG43069	cad	TF	3991961	3991971
Paris (France)	Jheh-3-CG43069	cad	TF	3991368	3991378
Paris (France)	Jheh-3-CG43069	cad	TF	3991686	3991696
Paris (France)	Jheh-3-CG43069	cad	TF	3991931	3991941
Paris (France)	Jheh-3-CG43069	Deaf1	TF	3991177	3991182
Paris (France)	Jheh-3-CG43069	hb	TF	3991367	3991376
Paris (France)	Jheh-3-CG43069	Ubx	TF	3991408	3991415
Paris (France)	Jheh-3-CG43069	Ubx	TF	3991412	3991419
Paris (France)	Jheh-3-CG43069	Ubx	TF	3991346	3991353
Paris (France)	Jheh-3-CG43069	Ubx	TF	3991410	3991417
Paris (France)	Jheh-3-CG43069	Ubx	TF	3991414	3991421

59

60

61