Tirant Stealthily Invaded Natural *Drosophila melanogaster* Populations during the Last Century

Florian Schwarz [iD],[1,2] Filip Wierzbicki,[1,2] Kirsten-André Senti,[†,1] and Robert Kofler [iD]*[,1]

[1]Institut für Populationsgenetik, Vetmeduni Vienna, Vienna, Austria
[2]Vienna Graduate School of Population Genetics, Vetmeduni Vienna, Vienna, Austria
[†]Present address: Institute of Molecular Biotechnology of the Austrian Academy of Sciences, Vienna Biocenter, Vienna, Austria
*Corresponding author: E-mail: rokofler@gmail.com.
Associate editor: John True

Abstract

It was long thought that solely three different transposable elements (TEs)—the I-element, the P-element, and hobo—invaded natural *Drosophila melanogaster* populations within the last century. By sequencing the "living fossils" of *Drosophila* research, that is, *D. melanogaster* strains sampled from natural populations at different time points, we show that a fourth TE, Tirant, invaded *D. melanogaster* populations during the past century. Tirant likely spread in *D. melanogaster* populations around 1938, followed by the I-element, hobo, and, lastly, the P-element. In addition to the recent insertions of the canonical Tirant, *D. melanogaster* strains harbor degraded Tirant sequences in the heterochromatin which are likely due to an ancient invasion, likely predating the split of *D. melanogaster* and *D. simulans*. These degraded insertions produce distinct piRNAs that were unable to prevent the novel Tirant invasion. In contrast to the I-element, P-element, and hobo, we did not find that Tirant induces any hybrid dysgenesis symptoms. This absence of apparent phenotypic effects may explain the late discovery of the Tirant invasion. Recent Tirant insertions were found in all investigated natural populations. Populations from Tasmania carry distinct Tirant sequences, likely due to a founder effect. By investigating the TE composition of natural populations and strains sampled at different time points, insertion site polymorphisms, piRNAs, and phenotypic effects, we provide a comprehensive study of a natural TE invasion.

Key words: transposable elements, *Drosophila melanogaster*, transposon invasions, next-generation sequencing, Tirant, P-element, I-element, hobo.

Introduction

Transposable elements (TEs) are DNA sequences that multiply within host genomes, even if this activity is deleterious to hosts (Doolittle and Sapienza 1980; Orgel and Crick 1980; Hickey 1982; Wicker et al. 2007). To enhance their rate of transmission into the next generation, TEs need to infect the germ cells. Although most TEs achieve this by being active in the germline, some LTR retrotransposons generate virus-like particles in the somatic follicle cells surrounding the germline, which may infect the germ cells (Song et al. 1997; Blumenstiel 2011; Goodier 2016; Moon et al. 2018; Wang et al. 2018). Since many TE insertions are deleterious, host organisms evolved elaborate defense mechanisms against TEs (Brennecke et al. 2007; Marí-Ordóñez et al. 2013; Yang et al. 2017). In *Drosophila melanogaster*, the defense against TEs is based on piRNAs (PIWI-interacting RNAs), that is, small RNAs with a size between 23–29nt, that repress TE activity at the transcriptional and the posttranscriptional level (Brennecke et al. 2007; Gunawardane et al. 2007; Sienski et al. 2012; Le Thomas et al. 2013). piRNAs are derived from distinct genomic loci termed piRNA clusters (Brennecke et al. 2007). Different piRNA pathways are active in the germline and in the follicle cells surrounding the germline (Li, Vagin, et al. 2009; Malone et al. 2009), where solely the germline pathway depends on maternally transmitted piRNAs for efficient silencing of TEs (Le Thomas et al. 2014).

One option to escape the host defense is to infect a novel species. Many TEs cross species boundaries, for example, due to horizontal transfer (HT) from one host species to another, and trigger invasions in naive species not having the TE (Mizrokhi and Mazo 1990; Maruyama and Hartl 1991; Lohe et al. 1995; Terzian et al. 2000; Sánchez-Gracia et al. 2005; Loreto et al. 2008; Kofler, Hill, et al. 2015; Peccoud et al. 2017). A striking example for a high frequency of TE invasions can be seen in *D. melanogaster*, which was invaded by at least three different TE families within the last century: the I-element, hobo, and the P-element (Kidwell 1983; Anxolabéhère et al. 1988; Periquet et al. 1989; Daniels, Chovnick, et al. 1990; Daniels, Peterson, et al. 1990; Bucheton et al. 1992; Bonnivard et al. 2000). All of these three TEs actively replicate only in the germline and induce some phenotypic effects, the hybrid dysgenesis (HD) symptoms, which historically led to the discovery of the recent TE invasions in *D. melanogaster* (Bingham et al. 1982; Calvi and Gelbart 1994; Biémont 2010; Moon et al. 2018; Wang et al. 2018). An important hallmark of these HD symptoms is that the direction of

Open Access

Article

crosses between two strains is important. The offspring of crosses between males carrying a genomic factor (the TE) and females not carrying this factor frequently show various symptoms (e.g., atrophic ovaries) whereas the offspring of the reciprocal crosses is usually free of symptoms (Bucheton et al. 1976; Kidwell et al. 1977; Blackman et al. 1987; Yannopoulos et al. 1987). Hence, hybrid dysgenesis has a cytoplasmic as well as a genomic component.

Although TEs were quickly identified as the responsible genomic factor, the cytoplasmic component, the maternally transmitted piRNAs, was discovered much later (Bingham et al. 1982; Brennecke et al. 2008). It was realized that the presence of an HD-inducing TE in a strain mostly depends on the sampling date of a strain, where more recently sampled strains frequently carry the TE while old strains, sampled before the invasion, do not. It was thus suggested that the HD-inducing TEs recently invaded *D. melanogaster* populations (Kidwell 1983; Periquet et al. 1994). These invasions were probably triggered by HT events, where the P-element was likely acquired from *D. willistoni* and the I-element as well as hobo possibly from *D. simulans* (or another species from the *simulans* clade) (Daniels, Chovnick, et al. 1990; Daniels, Peterson, et al. 1990; Simmons 1992; Loreto et al. 2008; Blumenstiel 2019). However, even the old strains carried short and highly degraded (probably inactive) fragments of the I-element and hobo, mostly in the heterochromatin (Bucheton et al. 1984, 1986, 1992; Daniels, Chovnick, et al. 1990). Hence, the I-element and hobo likely invaded *D. melanogaster* populations at least twice. Solely the P-element does not have substantial similarity to sequences in the *D. melanogaster* genome, which suggests that the P-element invaded *D. melanogaster* populations for the first time. *Drosophila melanogaster* strains sampled at different time points, previously labeled as the "living fossils" of *Drosophila* research (Bucheton et al. 1992), were not only used to discover the three recent TE invasions but also to estimate the timing of the invasions: the I-element invasion occurred presumably between 1930 and 1950, the hobo invasion around 1955 and the P-element invasion between 1950 and 1980 (Kidwell 1983; Anxolabéhère et al. 1988; Periquet et al. 1989).

By sequencing these "living fossils," we discovered that an additional transposon, Tirant, invaded *D. melanogaster* populations within the last century. Previous work showed that Tirant is an LTR retrotransposon and a member of the *Ty3/Gypsy* superfamily (Moltó et al. 1996; Viggiano et al. 1997; Cañizares et al. 2000; Terzian et al. 2001). It encodes an envelope protein and completes the retroviral cycle in the closely related *D. simulans* (Lemeunier et al. 1976; Marsano et al. 2000; Akkouche et al. 2012). In contrast to the P-element, hobo, and the I-element, which are active in the germline, Tirant was classified as an intermediate TE based on the amount of maternally transmitted piRNAs, that is, Tirant is likely expressed and targeted in both the germline and in somatic follicle cells (Malone et al. 2009). In agreement with this, Tirant activity was reported in both tissues (Akkouche et al. 2012). Furthermore, knockdowns of components of the germline as well as the somatic piRNA pathway, result in a reduction of Tirant piRNAs (Nefedova et al. 2012; Czech et al.

2013; Rozhkov et al. 2013; Barckmann et al. 2018). Generally, intermediate TEs are little understood. However, for Tirant in particular, peculiarities in the regulation were noted (Akkouche et al. 2013; Parhad et al. 2017; Wang et al. 2020). For example, in some backgrounds Tirant may be upregulated independent of piRNAs (Parhad et al. 2017).

Fablet et al. (2007) suggested that Tirant is an ancient TE that is largely vertically transmitted in the *D. melanogaster* species subgroup. Analyses of the reference genome of *D. melanogaster* revealed the presence of degraded Tirant insertions in the heterochromatin and full-length insertions in the euchromatin (Bowen and McDonald 2001; Mugnier et al. 2008). The heterochromatic insertions are likely ancient, possibly predating the split of *D. melanogaster* and *D. simulans*, whereas the euchromatic insertions are likely more recent (<16,000–200,000 years) (Bowen and McDonald 2001; Bergman and Bensasson 2007; Mugnier et al. 2008). This raises the question on how this uneven age distribution of Tirant insertions evolved.

Here, we show that full-length (canonical) Tirant sequences are absent from laboratory strains sampled before 1938 but present in strains sampled after 1938. We thus suggest that the canonical Tirant invaded natural *D. melanogaster* populations between 1930 and 1950, possibly following an HT from *D. simulans*. This invasion constitutes a second wave of activity, with degraded heterochromatic fragments being the remnants of an ancient Tirant invasion, possibly in the ancestor of the *D. melanogaster* species subgroup. Tirant is thus the fourth TE to invade *D. melanogaster* populations within the last century. Based on a consistent approach (i.e., the same method and strains) for all four TEs, we estimate that Tirant invaded *D. melanogaster* populations first, followed by the I-element, hobo and, finally, the P-element. Recent Tirant insertions were found in all investigated natural populations, where populations from Tasmania carry distinct Tirant sequences, likely due to a founder effect.

Although all strains carry piRNAs complementary to the degraded Tirant insertions solely recently invaded strains carry piRNAs complementary to the canonical Tirant. We thus suggest that piRNAs complementary to heterochromatic insertions were too diverged to prevent the spread of the canonical Tirant. Finally, we did not find apparent HD symptoms induced by Tirant, which may account for the late discovery of the Tirant invasion. By investigating the TE composition (i.e., abundance of TEs and frequency of internal deletions and SNPs) of natural populations and strains sampled at different time points, insertion site polymorphisms, piRNAs, and phenotypic effects, we provide a comprehensive study of a natural TE invasion.

Results

Canonical Tirant Insertions Are Present in Iso-1 but Not in Canton-S

Given the striking accumulation of TE invasions within the last century (Kidwell 1983; Anxolabéhère et al. 1988; Periquet et al. 1989; Daniels, Chovnick, et al. 1990; Daniels, Peterson, et al. 1990; Bucheton et al. 1992; Bonnivard et al. 2000), we

speculated that additional, hitherto undetected TEs, may have recently invaded *D. melanogaster* populations.

To test this hypothesis, we compared the abundance of TEs between one of the oldest available *D. melanogaster* laboratory strains, Canton-S (collected by C. Bridges in 1935; Lindsley and Grell 1968) and the reference strain, Iso-1 (fig. 1A; Brizuela et al. 1994). We aligned publicly available short-read data from these strains to the consensus sequences of TEs in *D. melanogaster* (Quesneville et al. 2005) and estimated the normalized abundance (reads per million) of the TEs in these two strains with our novel tool DeviaTE (Weilguny and Kofler 2019). Apart from the telomeric TEs (TART-A, TART-B, and TAHRE) which show distinct evolutionary dynamics (Pardue and DeBaryshe 2011; Saint-Leandre and Levine 2020), the most striking difference between the two strains was due to the LTR retrotransposon Tirant (fig. 1A). As expected, hobo and the I-element, two TEs that invaded *D. melanogaster* recently, are more abundant in the Iso-1 strain than in the older Canton-S strain (fig. 1A). The P-element is not present in both strains. To further investigate the abundance of Tirant in the two strains, we calculated the coverage of reads along the Tirant sequence with DeviaTE (fig. 1B; Weilguny and Kofler 2019). We observed striking coverage differences between Canton-S and Iso-1 over the entire sequence of Tirant (fig. 1B; average normalized coverage; Iso-1 = 20.9, Canton-S = 0.86). Only few highly diverged reads aligned to Tirant in Canton-S (fig. 1B). In addition to these diverged reads, many reads with a high similarity to the consensus sequence of Tirant aligned in Iso-1 (fig. 1B). We refer to Tirant sequences with a high similarity to the consensus sequence as "canonical" Tirant. To identify the genomic location of the canonical and the diverged Tirant sequences, we annotated TEs in publicly available assemblies of Canton-S (based on Oxford Nanopore long-read data) and Iso-1 (i.e., the reference genome) with RepeatMasker (fig. 1C; Hoskins et al. 2015; Wierzbicki et al. 2020). Both assemblies are of high quality and suitable for genomic analysis of TEs (Wierzbicki et al. 2020). In Canton-S, only highly fragmented and diverged Tirant sequences were found close to the centromeres (fig. 1C and supplementary fig. 1, Supplementary Material online). In addition to these diverged Tirant sequences, Iso-1 carries several canonical Tirant insertions on each chromosome arm (fig. 1C). This genomic distribution of Tirant, that is, degraded Tirant fragments in the heterochromatin and canonical insertions in the euchromatin of *D. melanogaster*, was also noted in previous studies (Marsano et al. 2000; Mugnier et al. 2008). The absence of canonical Tirant insertions in euchromatin is also found in an independent assembly of Canton-S which is based on PacBio reads (supplementary fig. 2, Supplementary Material online; Chakraborty et al. 2019). It was proposed that the degraded Tirant insertions located in heterochromatin are ancient and likely vertically inherited from the ancestor of the *D. melanogaster* species subgroup (Moltó et al. 1996; Fablet et al. 2007; Mugnier et al. 2008). It was further proposed that canonical insertions in Iso-1 are of more recent origin (i.e., < 16,000–200,000 years (Bowen and McDonald 2001; Bergman and Bensasson 2007; Lerat et al. 2011; Rahman

et al. 2015). We thus speculated that the canonical insertions of Tirant may have recently been active, whereas the degraded insertions in the heterochromatic may be inactive for some time (see also, Mugnier et al. 2008; Fablet et al. 2009). If this is true, canonical insertions ought to segregate at low frequency in natural populations, whereas the degraded insertions should mostly be fixed. To test this hypothesis, we estimated the population frequencies of the canonical and the degraded Tirant insertions in a natural *D. melanogaster* population from France (Viltain) (Kapun et al. 2020) with PoPoolationTE2 (Kofler et al. 2016). Indeed, most canonical Tirant insertions segregate at a low population frequency ($f = 0.063$) in the euchromatin, whereas most degraded insertions are in the heterochromatin and segregate at significantly higher frequencies ($f = 0.73$; Wilcoxon rank sum test $P < 2.2e{-}16$; supplementary fig. 3, Supplementary Material online). Due to relaxed purifying selection in low-recombining regions (Eanes et al. 1992; Sniegowski and Charlesworth 1994; Bartolomé et al. 2002; Petrov et al. 2011; Kofler et al. 2012), degraded Tirant insertions may have accumulated in the heterochromatin. Taken together, we hypothesize that Tirant invaded natural *D. melanogaster* populations in at least two waves of activity: an ancient wave, possibly predating the split of *D. melanogaster* and *D. simulans*, and a recent wave after Canton-S was sampled.

Canonical Tirant Invaded *D. melanogaster* Populations between 1930 and 1950

If Tirant invaded natural *D. melanogaster* populations recently, old strains should only have a few highly degraded Tirant sequences (similar to Canton-S), whereas more recently collected strains should have many insertions with a high similarity to the consensus sequence of Tirant (i.e., canonical Tirant insertions). To test this, we sequenced 12 of the oldest available *D. melanogaster* strains (sampled between 1920 and 1970; fig. 2; supplementary table 1, Supplementary Material online). Additionally, we included publicly available data of 15 different *D. melanogaster* strains into the analyses (fig. 2A and supplementary table 1, Supplementary Material online). The reads were mapped to the consensus sequences of TEs in *Drosophila* and the TE abundance was assessed with DeviaTE (supplementary fig. 4, Supplementary Material online; Weilguny and Kofler 2019).

Strikingly, six out of seven strains sampled before or in 1938 solely contained degraded Tirant sequences (supplementary table 1 and fig. 4, Supplementary Material online). The first strain carrying canonical Tirant sequences (Urbana-S) was collected around 1938. All 16 strains collected around or after 1950 carried canonical Tirant sequences (supplementary table 1, Supplementary Material online). Estimates of the TE copy numbers support these observations (fig. 2A). To obtain estimates of the TE abundance independent of DeviaTE, we also computed the normalized number of reads mapping to each TE (rpm; reads per million). These data also support the sudden increase in reads mapping to Tirant in strains sampled after 1938 (supplementary table 2, Supplementary Material online). We note that the raw abundance of reads mapping
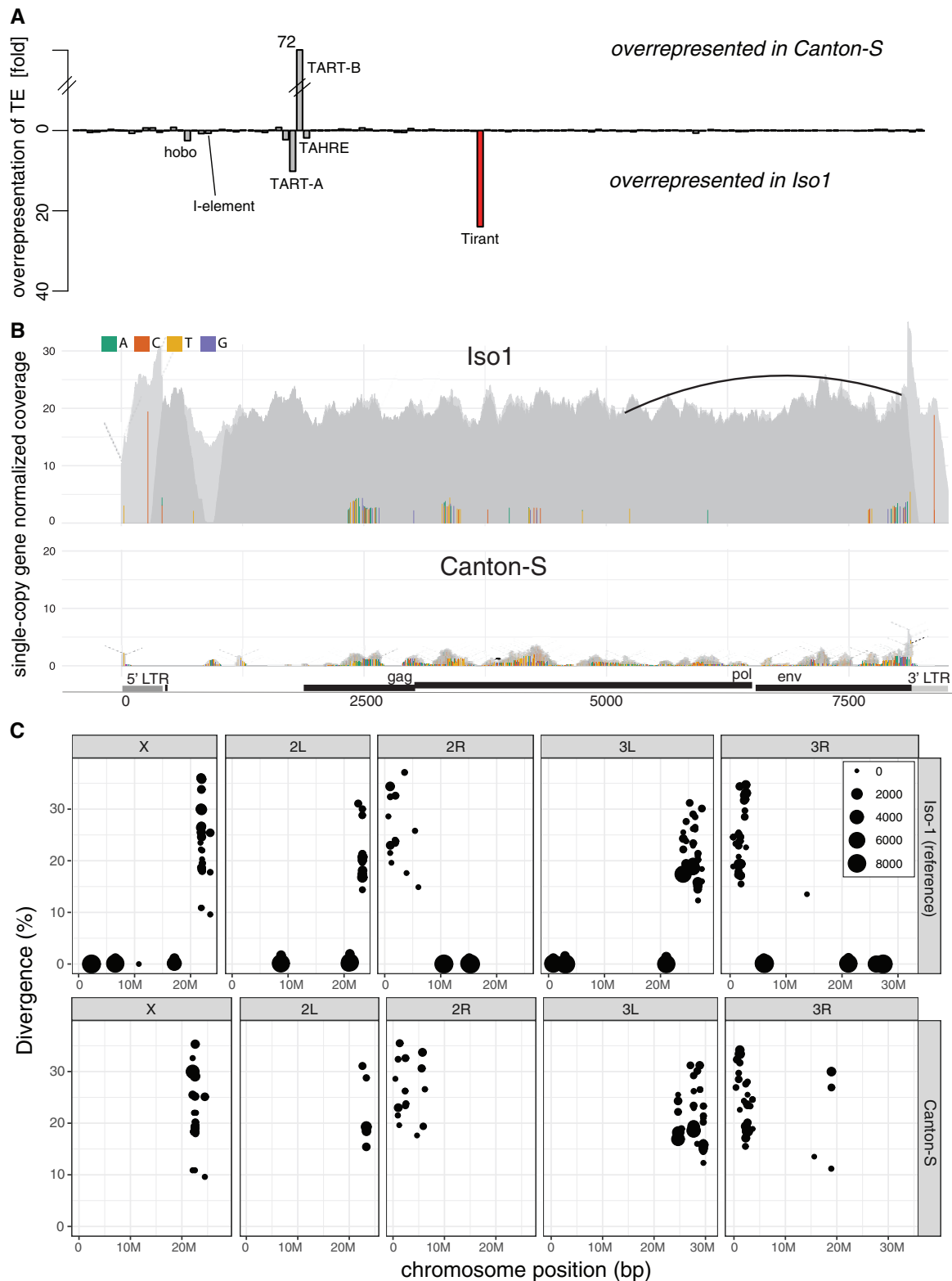
FIG. 1. Canonical Tirant insertions are present in Iso-1 but not Canton-S. (*A*) Differences in TE content between Canton-S and Iso-1. For each TE family (*x* axis), we show the fold-difference in the number of reads mapping to a TE (*y* axis) between the two strains. Note that reads mapping to Tirant (red) are overrepresented in Iso-1. (*B*) Abundance and diversity of Tirant in Iso-1 and Canton-S. Short reads were aligned to the consensus sequence of Tirant and visualized with DeviaTE. The coverage of Tirant was normalized to the coverage of single-copy genes. Single-nucleotide polymorphisms (SNPs) and small internal deletions (indels) are shown as colored lines. Large internal deletions are shown as black arcs (the frequency of the shown deletion is ≈5%). Coverage based on uniquely and ambiguously mapped reads is shown in dark and light gray, respectively. Note that solely a few, highly degraded copies of Tirant are present in Canton-S. (*C*) Overview of Tirant insertions in the genomes of Iso-1 and Canton-S. For each Tirant insertion, we show the position in the assembly, the length (size of dot), and the similarity to the consensus sequence (divergence).
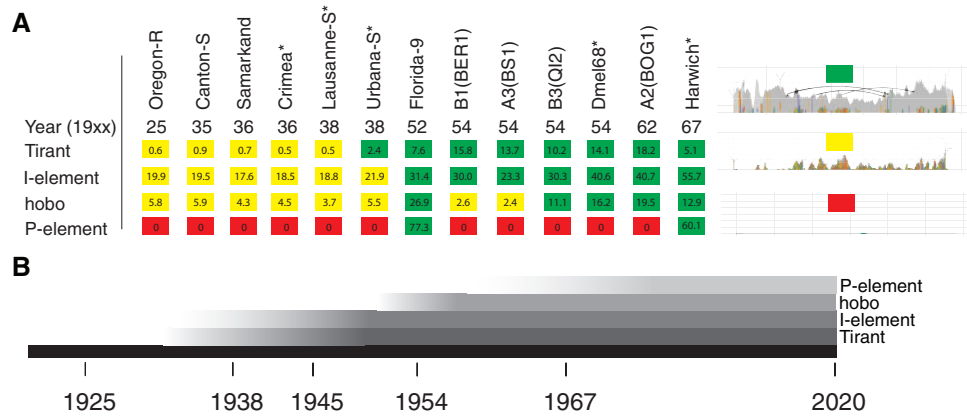
	Oregon-R	Canton-S	Samarkand	Crimea *	Lausanne-S*	Urbana-S*	Florida-9	B1(BER1)	A3(BS1)	B3(QI2)	Dmel68*	A2(BOG1)	Harwich*
Year (19xx)	25	35	36	36	38	38	52	54	54	54	54	62	67
Tirant	0.6	0.9	0.7	0.5	0.5	2.4	7.6	15.8	13.7	10.2	14.1	18.2	5.1
I-element	19.9	19.5	17.6	18.5	18.8	21.9	31.4	30.0	23.3	30.3	40.6	40.7	55.7
hobo	5.8	5.9	4.3	4.5	3.7	5.5	26.9	2.6	2.4	11.1	16.2	19.5	12.9
P-element	0	0	0	0	0	0	77.3	0	0	0	0	0	60.1

Fig. 2. History of recent TE invasions in natural *Drosophila melanogaster* populations. (*A*) Overview of Tirant, I-element, hobo, and P-element sequences in some *D. melanogaster* strains. For an overview of these TEs in all investigated strains, see supplementary table 1, Supplementary Material online. The year corresponds to the sampling date of the strain. Strains sequenced in this work are marked by a star (*). For each family, we classified the TE content into three distinct categories: red, absence of any TE sequence; yellow, solely degraded TE sequences are present; green, nondegraded sequences, with a high similarity to the consensus sequence are present; Numbers represent estimates of TE copy numbers per haploid genome obtained with DeviaTE. The abundance of degraded copies may be underestimated as copy-number estimates are based on the average coverage of the consensus sequence. The right panel shows an example for each of the three categories (similar to fig. 1*B*). (*B*) Timeline showing the estimated invasion history of Tirant, the I-element, hobo, and the P-element.

to a TE is highly correlated with the estimates of TE abundance obtained with DeviaTE (supplementary fig. 5, Supplementary Material online). Our results thus suggest that the canonical Tirant invaded *D. melanogaster* populations between 1938 and 1950 (fig. 2). Since we were interested in the timing of the Tirant invasion relative to the other three TEs that recently invaded *D. melanogaster* populations, we also investigated the abundance and diversity of the I-element, hobo, and the P-element in these strains (supplementary table 1 and figs. 6–8, Supplementary Material online). Our data suggest that Tirant invaded natural *D. melanogaster* populations just before the I-element, followed by hobo and, lastly, by the P-element (supplementary tables 1 and 2, Supplementary Material online and fig. 2*B*).

Canonical Tirant Insertions Are Found in Worldwide Populations of *D. melanogaster* and Populations from Tasmania Carry Distinct Tirant Variants

To further investigate the Tirant composition among strains, we performed a PCA based on the allele frequencies of Tirant single-nucleotide polymorphism (SNPs) (fig. 3). Note that our usage of the term SNP is not strictly identical to the common usage describing allelic variants at a single locus. Here, a SNP describes a variant among dispersed Tirant copies. Our allele frequency estimates thus reflect the Tirant composition within a particular strain (e.g., if 14 Tirant insertions in a given strain carry an "A" at some site and 6 a "T," the frequency of "A" at this site is 0.7). In addition to the above-mentioned strains (supplementary table 1, Supplementary Material online), we also analyzed the Tirant content of natural populations. To do this, we relied on the global diversity lines (GDL), that is, several *D. melanogaster* strains sampled after 1988 (Begun and Aquadro 1995) from five different continents (Africa—Zimbabwe, Asia—Beijing, Australia—Tasmania, Europe—Netherlands, America—Ithaca; Grenier et al. 2015).

Old strains, collected before 1938, formed a distinct group (fig. 3), supporting our view that they carry distinct Tirant sequences. By contrast, most strains collected after 1938 and the majority of the GDLs group into one large cluster (fig. 3). All GDL strains thus carry nondegraded Tirant sequences. This observation also holds when additional, recently collected *D. melanogaster* strains are analyzed (e.g., DGRP, DrosEU, DrosRTEC; supplementary fig. 9, Supplementary Material online; Mackay et al. 2012; Bergland et al. 2014; Lack et al. 2015; Machado et al. 2019; Kapun et al. 2020). Old strains also form a distinct group in an unrooted tree computed from pairwise F_{ST} values based on the frequency of Tirant SNPs (supplementary fig. 10, Supplementary Material online). Our data thus suggests that Tirant invaded most worldwide *D. melanogaster* populations. The reference strain Iso-1 is distant to the large cluster (fig. 3). Closer inspection revealed that Tirant insertions from natural populations carry eight SNPs that are not found in the reference strain (supplementary fig. 11 and table 3, Supplementary Material online). Interestingly, also strains collected from Tasmania (Australia) formed a distinct group (fig. 3 and supplementary fig. 10, Supplementary Material online). We hypothesized that this is due to multiple SNPs having markedly different allele frequencies in Tasmanian populations than in populations from other geographic locations (supplementary fig. 12 and table 4, Supplementary Material online). Indeed, when excluding those SNPs from the PCA, strains from Tasmania clustered with strains sampled from the other locations (supplementary fig. 13, Supplementary Material online). For hobo and the I-element, Tasmanian populations did not form a separate cluster (supplementary fig. 14, Supplementary Material online; the P-element is absent in many samples, hence allele frequencies could not be calculated). This raises the question of what processes could be responsible for such striking differences in the Tirant composition among natural
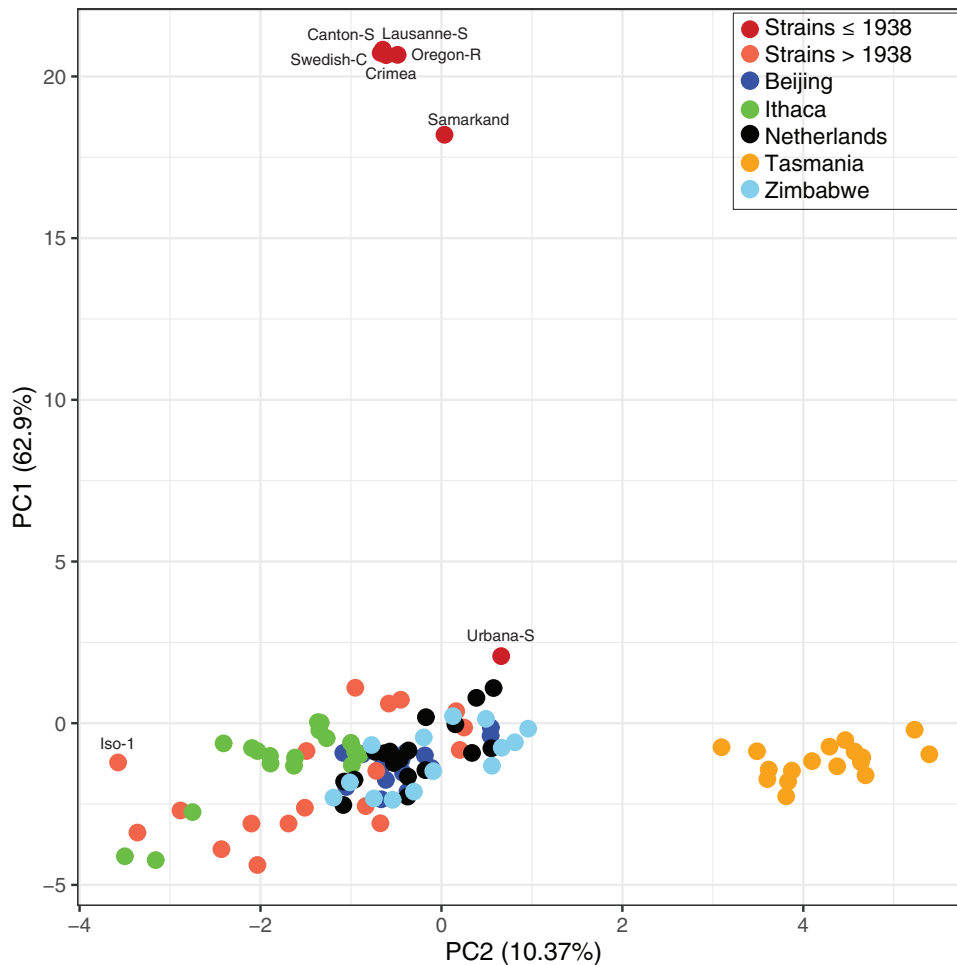
FIG. 3. PCA based on the allele frequencies of Tirant SNPs in different *Drosophila melanogaster* strains. In addition to previously described *D. melanogaster* strains, the Global Diversity Lines (GDL) were analyzed. Note that the strains without canonical Tirant insertions as well as populations from Tasmania form distinct groups.

populations. We suggest that the Tirant invasion in Tasmania was subject to a founder effect, where flies carrying some rare variants of Tirant migrated to Tasmania, thereby triggering the spread of these rare Tirant variants in Tasmanian populations. Similarly, the strains used for generating Iso-1 may have carried rare Tirant variants that multiplied in these lines after they were sampled. In agreement with this, most Iso-1 specific SNPs segregate at low frequency in some *D. melanogaster* populations from Europe and North America (supplementary fig. 11, Supplementary Material online).

In summary, we conclude that Tirant invaded all investigated worldwide populations of *D. melanogaster* during the past century. Furthermore, founder effects may be important components of TE invasions, since they may lead to a geographically heterogeneous TE composition.

The Canonical Tirant Is Silenced by a piRNA-Based Defense Mechanism

If Tirant recently invaded *D. melanogaster* populations, we expect to see differences in the composition of piRNAs between strains sampled before and after the invasion. Strains invaded by Tirant, such as Iso-1, should have established a

functional defense against the TE and thus generate large amounts of piRNAs complementary to canonical Tirant. By contrast, naive strains, such as Canton-S, should have few canonical Tirant piRNAs. To test this, we sequenced piRNAs from the ovaries of both strains. Indeed, piRNAs against canonical Tirant were highly abundant in Iso-1 but not in Canton-S (fig. 4A and D). Compared with the piRNA abundance of other TE families in *D. melanogaster*, Tirant piRNAs rank among the most abundant in Iso-1 but the least abundant in Canton-S (fig. 4A). Both sense and antisense piRNAs are distributed over the entire sequence of Tirant in Iso-1 (fig. 4B). TEs that are silenced in the germline by dual-strand clusters show a characteristic 10 nt overlap between sense and antisense piRNAs, that is, the ping-pong signature (Brennecke et al. 2007; Malone et al. 2009). Tirant has a pronounced ping-pong signature in Iso-1 but not in Canton-S (fig. 4C), consistent with Tirant being silenced in the germline (likely in addition to the soma) (Malone et al. 2009). Finally, we wondered whether the ancient Tirant invasion, responsible for the degraded Tirant fragments in the heterochromatin, led to piRNAs against Tirant. Both Iso-1 and Canton-S, carry piRNAs complementary to the degraded Tirant fragments (6252.0 ppm in Canton-S and
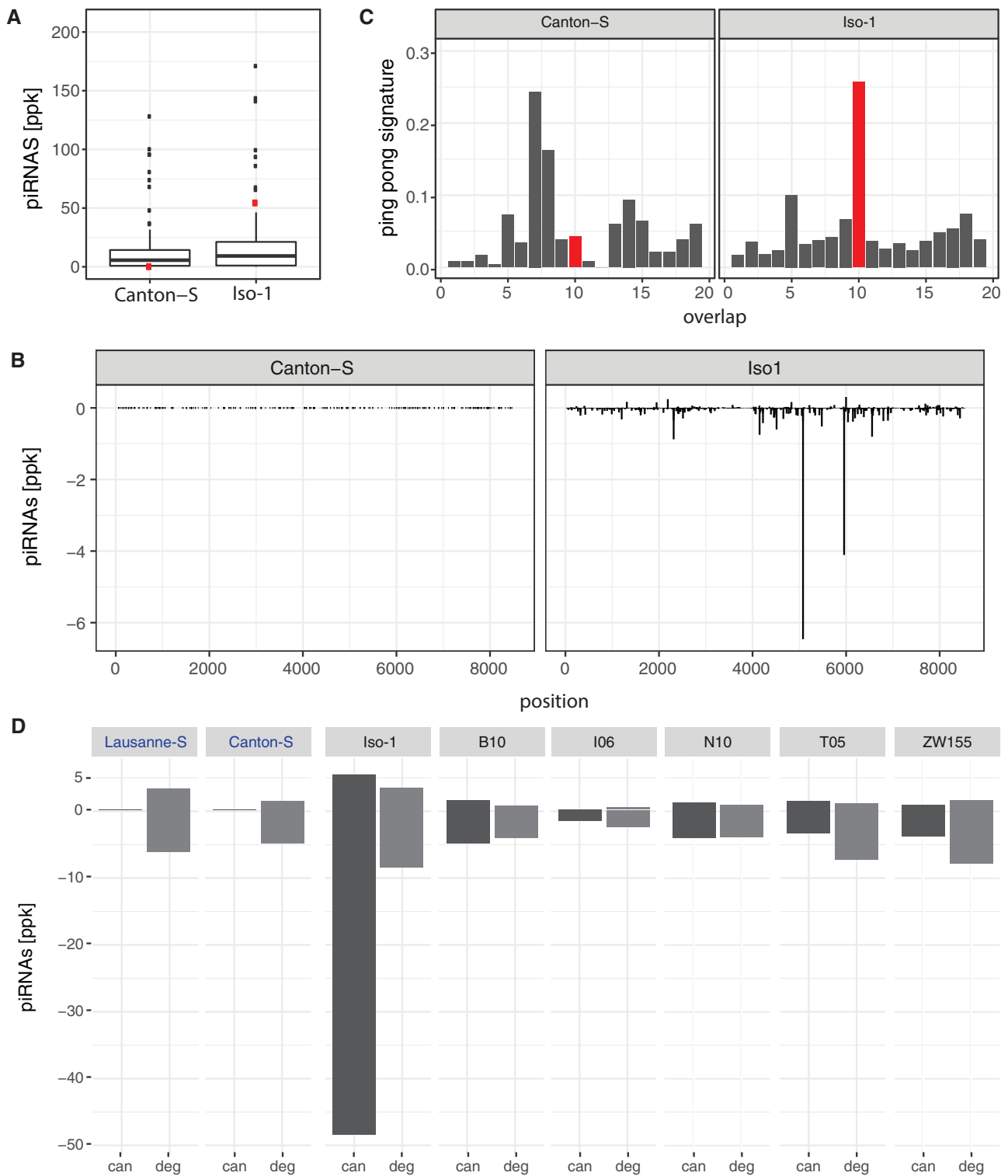
FIG. 4. Tirant piRNAs in strains with (e.g., Iso-1) and without (e.g., Canton-S) canonical Tirant insertions. (A) Abundance of canonical Tirant piRNAs (red) compared with piRNA of the other TEs of *Drosophila melanogaster*. (B) Abundance of piRNAs along the canonical sequence of Tirant. (C) Ping-pong signature for canonical Tirant piRNAs. A pronounced peak at position 10 (red) suggests secondary amplification of piRNAs by the ping-pong cycle. (D) Abundance of piRNAs complementary to canonical (can; dark gray) and degraded (deg; light gray) Tirant sequences in laboratory strains (Lausanne-S, Canton-S, and Iso-1) and GDL lines (B10, Beijing; I06, Ithaca; N10, Netherlands; T05, Tasmania; ZW155, Zimbabwe; Grenier et al. 2015). The names of two strains not having canonical Tirant insertions are shown in blue. Sense piRNAs are on the positive *y* axis and antisense piRNAs on the negative *y* axis (B and D). ppk, piRNAs per 1000 miRNAs.

11886.0 ppm in Iso-1; fig. 4D). An analysis of the piRNA content of additional strains (Lausanne-S and GDL lines; Luo et al. 2020) confirms that all investigated strains carry piRNAs complementary to the degraded Tirant whereas only strains with canonical Tirant insertions carry piRNAs complementary to the canonical Tirant (fig. 4D). This raises the question why these piRNAs of the degraded Tirant were unable to prevent the invasion of the canonical Tirant. Previous works suggest that piRNAs need to match over the bulk of a sequence with a sequence divergence of less than 10% for efficient silencing of the target sequence (Post et al. 2014; Kotov et al. 2019). Heterochromatic Tirant sequences, however, are about 10–30% diverged from the canonical Tirant (supplementary fig. 1, Supplementary Material online). The high divergence can be found over the entire sequence of these Tirant fragments (supplementary fig. 15, Supplementary Material online). Consequently very few of the degraded piRNAs match to the canonical Tirant with a sequence divergence of less than 10% (supplementary fig. 16, Supplementary Material online).

We conclude that a piRNA-based defense mechanism against the canonical Tirant is present in all strains carrying canonical Tirant insertions but absent in strains solely having heterochromatic Tirant insertions. Although piRNAs derived from these heterochromatic Tirant fragments are present in all strains, these piRNAs were likely too diverged to silence the canonical Tirant and therefore could not prevent its recent invasion.

No Apparent Hybrid Dysgenesis Symptoms Can Be Found for Tirant

The other three TEs that invaded *D. melanogaster* populations within the last 100 years (I-element, hobo, P-element) caused some hybrid dysgenesis (HD) symptoms. To test whether Tirant also induces HD symptoms, we performed crosses between strains having recent Tirant insertions (Urbana-S and Hikone-R) and strains not having such insertions (Lausanne-S and Canton-S). All strains do not have recent P-element, I-element, and hobo insertions, which rules out interference by the other HD systems (fig. 2A and supplementary table 1, Supplementary Material online). We investigated the fraction of dysgenic ovaries in the F1 generation, a trait influenced by P-element and hobo mobilization (Kidwell et al. 1977; Blackman et al. 1987; Yannopoulos et al. 1987), and the fraction of hatched F2 embryos, a trait influenced by I-element mobilization (Bucheton et al. 1976). We performed all crosses at several temperatures (supplementary fig. 17A and B, Supplementary Material online), as temperature frequently has a strong influence on the extent of HD symptoms (Kidwell et al. 1977; Bucheton 1979; Kidwell and Novy 1979; Serrato-Capuchina et al. 2020). We did not find any significant differences in the number of dysgenic ovaries nor in the number of hatched eggs between the reciprocal crosses (supplementary fig. 17A and B and table 5, Supplementary Material online). As the number of paternally inherited TEs may influence the magnitude of HD (Serrato-Capuchina et al. 2020), we performed reciprocal crosses with the strain carrying the largest number

of canonical Tirant insertions, that is, Iso-1, and strains not having canonical Tirant insertions (Lausanne-S and Crimea; supplementary fig. 1, Supplementary Material online). However, Iso-1 also carries I-element and hobo insertions (supplementary fig. 1, Supplementary Material online). Therefore, we performed crosses solely at 25 °C, a temperature where I-element HD is usually not observed (Bucheton et al. 1976). As strains inducing hobo HD are very rare (Pascual and Periquet 1991), there is solely a small chance that hobo activity will generate atrophic ovaries in this crosses. We again did not find any significant differences in the number of dysgenic ovaries nor in the number of hatched eggs among these crosses (supplementary fig. 17C and D and table 5, Supplementary Material online; which also rules out hobo HD).

We hypothesize that the absence of apparent HD symptoms may be one reason why the invasion of Tirant in natural *D. melanogaster* populations during the past century was not detected before.

Origin of the Canonical Tirant Invasion

Lastly, we aimed to shed light on the origin of the Tirant invasion. Since canonical Tirant insertions are mostly absent in strains collected before 1938, we reasoned that the recent Tirant invasion was likely triggered by HT (or an introgression). To identify the putative donor species, we investigated Tirant sequences in different *Drosophila* species. We first tested if Tirant sequences can be found in 11 sequenced *Drosophila* genomes (Drosophila 12 Genomes Consortium 2007). Solely members of the *D. melanogaster* species subgroup contained reads mapping to Tirant (supplementary fig. 18, Supplementary Material online; *D. melanogaster*, *D. simulans*, *D. erecta*, *D. yakuba*; in agreement with Fablet et al. [2007]). We also found that *D. simulans* is the only species that may carry full-length insertions of Tirant (apart from *D. melanogaster*) and that some Tirant insertions in *D. simulans* may have a high similarity to the consensus sequence of Tirant (supplementary fig. 18, Supplementary Material online). To further investigate the composition of Tirant in the *D. melanogaster* species subgroup, we obtained Illumina short-read data for several individuals from different species of this subgroup. In addition to *D. melanogaster*, *D. simulans*, *D. erecta*, and *D. yakuba*, we also obtained data for *D. sechellia*, *D. mauritiana*, and *D. teisseri* (supplementary table 6, Supplementary Material online). A PCA based on the allele frequencies of Tirant SNPs confirms that the Tirant composition of recently collected *D. melanogaster* strains (>1938) is most similar to *D. simulans* strains (supplementary fig. 19, Supplementary Material online). The high similarity of some Tirant sequences between *D. melanogaster* and *D. simulans* was noted before (Fablet et al. 2006; Lerat et al. 2011; Bargues and Lerat 2017). However, an analysis based on the allele frequencies confounds the two subfamilies of Tirant in these two species, for example, canonical Tirant insertions (Tirant-C in *D. simulans*) and degraded Tirant insertions (Tirant-S in *D. simulans*) (Fablet et al. 2006). Therefore, to further investigate whether some Tirant insertions of *D. simulans* could have triggered the canonical Tirant invasion

in *D. melanogaster*, we analyzed the Tirant content in a recent long-read based assembly of *D. simulans* (strain w^{XD1}; Chakraborty et al. 2020). Indeed, we found that *D. simulans* carries three full-length insertions that have a high similarity to the consensus sequence of Tirant (average divergence: 1.97%, 1.56%, 1.60%; supplementary table 7, Supplementary Material online). We concluded that HT from *D. simulans* may have triggered the invasion of the canonical Tirant in *D. melanogaster* populations.

Discussion

We show that the retrotransposon Tirant invaded most natural *D. melanogaster* populations between 1930 and 1950, possibly following HT from *D. simulans*. Tirant is thus the fourth TE that invaded *D. melanogaster* in the last century. We also provide the first comprehensive timeline of the recent TE invasions in *D. melanogaster* populations that is based on a consistent approach (i.e., the same method and strains). The canonical Tirant invaded natural *D. melanogaster* populations first followed by the I-element, hobo, and finally by the P-element. All investigated strains, including those lacking canonical Tirant insertions, carry highly degraded Tirant fragments, which likely stem from an ancient Tirant invasion predating the split of the *D. melanogaster* species subgroup (Fablet et al. 2007; Lerat et al. 2011). We demonstrate that piRNAs derived from canonical and diverged Tirant insertions can be clearly distinguished and suggest that piRNAs derived from degraded Tirant copies, which were present in all investigated strains, were unable to prevent the invasion of the canonical Tirant. We show that founder effects may be important components of TE invasions that may lead to a heterogeneous TE composition among populations. Finally, we did not find apparent HD symptoms among reciprocal crosses of strains with and without canonical Tirant insertions.

Our conclusion that Tirant recently invaded *D. melanogaster* is mainly based on the absence of canonical Tirant sequences in most strains collected before 1938 and their presence in strains collected after 1938. As an alternative explanation, most strains collected before 1938 could have lost the canonical Tirant sequences. It was, for example, proposed that non-African *D. simulans* populations lost canonical Tirant sequences (Fablet et al. 2006). But this alternative explanation seems unlikely as it requires the independent loss of canonical Tirant sequences in strains collected before 1938 but not in any strain collected after 1938. The low population frequency of euchromatic Tirant insertions (see also Kofler, Nolte, et al. 2015) and the high sequence similarity between the left and the right LTR of Tirant insertions (Bowen and McDonald 2001; Bergman and Bensasson 2007) are also in agreement with our hypothesis of a recent Tirant invasion. Our hypothesis of the recent Tirant invasion is also consistent with the interpretation of the data for the I-element, P-element, and hobo, where the absence of the (canonical) TE in old strains combined with the presence in young strains was taken as evidence for recent invasions of these elements

(Kidwell 1983; Daniels, Chovnick, et al. 1990; Daniels, Peterson, et al. 1990; Bucheton et al. 1992).

Our data suggest that Tirant was the first TE that invaded natural *D. melanogaster* populations in the last century. However, these results need to be interpreted with caution as 1) there is some uncertainty about the sampling time of the strains, 2) some strains may have been contaminated (e.g., the presence of the P-element in a strain collected around 1938 [Swedish-C] is likely due to mixing of strains during maintenance of stocks; supplementary table 1, Supplementary Material online), and 3) our strains are from different geographic regions, where some regions might have been invaded earlier than others. Nevertheless, our results are largely in agreement with previous works which suggested that the I-element invasion happened between 1930 and 1950, the hobo invasion around 1955 and the P-element invasion between 1950 and 1980 (Kidwell 1983; Anxolabéhère et al. 1988; Periquet et al. 1989).

We did not find evidence that Tirant induces HD symptoms. Also, a previous work in *D. simulans* did not report HD symptoms for Tirant despite Tirant being activated by reciprocal crosses (Akkouche et al. 2013). However, due to several reasons, more work will be necessary to show whether or not Tirant causes some HD symptoms. First, it is not clear what symptoms to look for. We investigated the fraction of dysgenic ovaries in the F1 and the fraction of hatched eggs (F2), two traits affected by HD from P-element, I-element, or hobo. However, it is feasible that Tirant activity leads to entirely different phenotypic effects, especially given that Tirant may be active in the germline and in the soma (Malone et al. 2009; Akkouche et al. 2013; Czech et al. 2013), and could thus affect both tissues. Second, it is not clear if intermediate TEs, such as Tirant, are able to induce HD. Different phenotypes among reciprocal crosses (i.e., HD) can solely be observed if maternally transmitted piRNAs (i.e., the cytoplasmic component of HD) are necessary to silence a TE (Brennecke et al. 2008). Maternally transmitted piRNAs initiate the ping-pong cycle and recruit silencing chromatin that is then bound by Rhino, which in turn defines the site of dual-strand clusters (Le Thomas et al. 2014). As both ping-pong and dual-strand clusters are solely observed in the germline piRNA pathway (Malone et al. 2009), it is thought that maternally deposited piRNAs are important for the germline pathway but not for the somatic piRNA pathway. Consequently, no HD symptoms are expected for TEs that are solely active in the soma. The I-element, hobo, and the P-element, three TEs that invaded *D. melanogaster* populations recently, were all active in the germline and induced HD symptoms (Bingham et al. 1982; Calvi and Gelbart 1994; Biémont 2010; Moon et al. 2018; Wang et al. 2018). However, the situation is entirely unclear for intermediate elements such as Tirant. Surprisingly, one study even suggested that maternally transmitted piRNAs are necessary to silence Tirant in the soma (Akkouche et al. 2013). The molecular mechanisms behind this influence of maternal piRNAs on the somatic piRNA pathway remain yet unclear. Third, the severity of HD symptoms frequently depends on multiple factors, such as temperature and the age of flies (Kidwell et al. 1977; Bucheton

1979; Kidwell and Novy 1979; Serrato-Capuchina et al. 2020). It is feasible that HD symptoms of Tirant can only be observed under certain conditions, and these conditions could differ substantially from the previously described HD systems. Fourth, previous studies noted marked differences in the ability to induce or repress HD among different strains (Kidwell et al. 1977, 1988; Anxolabéhère et al. 1988; Pascual and Periquet 1991; Srivastav et al. 2019). This could be mediated by differences in the number of paternally transmitted TEs (Srivastav and Kelleher 2017; Serrato-Capuchina et al. 2020), different variants of the TEs (Srivastav et al. 2019), and differences in the tolerance to TE activity among strains (Kelleher et al. 2018). The abundance of strains inducing HD may also vary among the HD systems. For example, strains inducing P-element HD are readily found whereas strains inducing hobo HD are rare (Kidwell 1983; Pascual and Periquet 1991). It is thus feasible that solely crosses of certain strains show HD symptoms of Tirant.

It is currently unclear how canonical Tirant sequences entered *D. melanogaster* populations. Possible explanations are HT or introgression from a related species (Silva et al. 2004; Sánchez-Gracia et al. 2005; Loreto et al. 2008; Bartolomé et al. 2009). In search for a possible donor species, we found that *D. simulans* carries some full-length Tirant insertions with a high similarity to canonical Tirant in *D. melanogaster* (supplementary table 7, Supplementary Material online). Out of the two Tirant subfamilies found in *D. simulans*, Tirant-C (nondegraded insertions) and Tirant-S (degraded insertions), Tirant-C insertions have been previously shown to be closely related to the canonical Tirant in *D. melanogaster* (Fablet et al. 2006; Lerat et al. 2011; Bargues and Lerat 2017). We thus suggest that HT of Tirant-C from *D. simulans* to *D. melanogaster* may have triggered the canonical Tirant invasion in *D. melanogaster*, in agreement with Lerat et al. (2011). Apart from this HT, Tirant is likely mostly vertically transmitted in the *D. melanogaster* species subgroup (Fablet et al. 2007). In agreement with this, a tree based on frequency of Tirant SNPs largely follows the species tree (supplementary fig. 20, Supplementary Material online). HT of TEs between *D. melanogaster* and *D. simulans* is plausible since both species are closely related (Lemeunier et al. 1976) and have largely overlapping habitats (Parsons and Stanley 1981), which generates ample opportunities for HT or introgressions. HT of TEs between these species was observed before in both directions. For example, Kofler, Hill, et al. (2015) suggested that *D. simulans* recently acquired the P-element from *D. melanogaster*. Conversely, hobo and the I-element in *D. melanogaster* were possibly acquired from *D. simulans* (Daniels, Chovnick, et al. 1990; Simmons 1992; Loreto et al. 2008).

We found that Tirant sequences from Tasmania (an island south of Australia) have a different composition than Tirant sequences from other locations (at least five SNPs have distinctly different frequencies; supplementary table 4, Supplementary Material online). We suggest that this may be due to a founder effect during the Tirant invasion, which led to the spread of rare Tirant variants in Tasmanian populations. We wondered whether the observed founder effect could be due to the recent colonization of Australia (Tasmania) by *D. melanogaster* (Bock and Parsons 1981). However, this seems unlikely as the colonization of Australia, and probably also of Tasmania, predates the Tirant invasion. *Drosophila melanogaster* was first spotted in Australia in 1894 and is known to rapidly spread into nearby areas (Bock and Parsons 1981; Keller 2007), whereas the Tirant invasion mostly happened between 1938 and 1950. Moreover, founder effects that occurred during the colonization of Tasmania should affect the entire genomic background of *D. melanogaster* and not just the Tirant sequences. Previous studies did not detect any signatures of bottlenecks for Tasmanian *D. melanogaster* populations (Agis and Schlötterer 2001; Grenier et al. 2015; Bergland et al. 2016; Arguello et al. 2019). We thus argue that the founder effect in Tasmania is specific to Tirant. Founder effects during TE invasions could be important, hitherto little considered, processes that may lead to geographically distinct TE variants.

We suggest that four different TEs invaded *D. melanogaster* populations within 40 years (between the 1930s and 1970s). Why did so many different TEs spread in *D. melanogaster* within such a short time? A possible explanation could be the recent habitat expansion of *D. melanogaster* into the Americas and Australia about 100–200 years ago (Bock and Parsons 1981; Vieira et al. 1999; Kofler, Nolte, et al. 2015). Habitat expansion may bring species into contact that did not coexist before in the same habitat. If these species carry different TE families, HT events between the species may trigger novel TE invasions. A classic example is the P-element in *D. melanogaster* which was likely acquired from *D. willistoni* after *D. melanogaster* entered the habitat of *D. willistoni* in South America (Engels 1992). The lag-time between colonization of the Americas and Australia (~100–200 years ago; Bock and Parsons 1981; Keller 2007) and the four different TE invasions (1930–1970) may be due to the stochasticity of HT events, a strong influence of drift in the early stages of TE invasions and the time required until a TE reaches an appreciable frequency (Ginzburg et al. 1984; Le Rouzic and Capy 2005). It will be interesting to see if such a high rate of novel TE invasions in *D. melanogaster* populations will be maintained over the next century. An absence of novel invasions would support our hypothesis that the habitat expansion triggered the four recent TE invasions in *D. melanogaster*.

Out of the four TEs that invaded *D. melanogaster* populations in the last century, the P-element is unique as it is the only TE that does not show substantial similarity to any sequence of the *D. melanogaster* genome. For the other three TEs—Tirant, the I-element, and hobo—many degraded insertions can be found (mostly in the heterochromatin) (Bucheton et al. 1984, 1986, 1992; Daniels, Chovnick, et al. 1990). Thus, three out of the four TEs probably invaded *D. melanogaster* populations at least twice. This raises the question of how multiple waves of invasions arise. Before a TE can trigger a novel invasion the TE needs to overcome the host defense (or the host defense may break down). For example, in mammals and invertebrates efficient silencing of a TE requires piRNAs that match with less than 10% sequence divergence over the bulk of the TE sequence (Post et al. 2014;

Kotov et al. 2019). A TE that diverged by more than 10% from the piRNA pool of the host (e.g., the canonical Tirant compared with the degraded Tirant sequences) could thus trigger a second wave of an invasion. The same consideration holds for other host defense mechanism that rely on sequence similarity to a TE, like small RNAs in plants or Kruppel-associated box zinc-finger proteins in mammals (Marí-Ordóñez et al. 2013; Yang et al. 2017). It is however an important open question whether sufficient sequence divergence could be acquired within a host species, where host defense mechanisms may coadapt with the TE, or whether HT to an intermediate host (e.g., a closely related species) is necessary to overcome the host defense.

Materials and Methods

Strains and Dating
The sequenced fly strains were obtained from the Bloomington Drosophila Stock Center (BDSC) (Crimea, Lausanne-S, Swedish-C, Urbana-S, Berlin-K, Hikone-R, Florida-9, Pi2, Harwich, Amherst-3) and the National Drosophila Species Stock Center (Dmel68). w1118 and wk were kindly provided by Silke Jensen. We additionally analyzed publicly available sequencing data of different *D. melanogaster* strains (King et al. 2012; Mackay et al. 2012; Bergland et al. 2014; Grenier et al. 2015; Lack et al. 2015; Jakšić et al. 2017; Machado et al. 2019; Kapun et al. 2020; Wierzbicki et al. 2020) (supplementary table 6, Supplementary Material online). The collection dates of the strains were obtained from different sources. If available, we used the collection dates from Lindsley and Grell (1968). Alternatively, we used the collection dates published in previous works (Black et al. 1987; Anxolabéhère et al. 1988; Galindo et al. 1995; Engels 2007; Ruebenbauer et al. 2008) or information from the National Drosophila Species Stock Center (drosophilaspecies.com) and FlyBase (flybase.org/reports/FBrf0222222.html, last accesssed December 15, 2020) (supplementary table 1, Supplementary Material online). For the strains w1118 and Urbana-S, we used the latest possible collection date: for w1118, we used the publication date of the first publication mentioning the strain and for Urbana-S, we used the year of the death of C. Bridges, who collected the strain (Lindsley and Grell 1968) (supplementary table 1, Supplementary Material online). The geographic origin was obtained from the same sources. For an overview of the used strains, the estimated collection date, and the source of the information, see supplementary table 1, Supplementary Material online. The Iso-1 strain was generated by crossing several laboratory strains, with largely unknown sampling dates (Brizuela et al. 1994). Therefore, we did not assign a sampling date to this strain. Additionally, we used publicly available data of different strains from *D. simulans*, *D. sechellia*, *D. mauritiana*, *D. yakuba*, *D. teisseri*, and *D. erecta* (*Drosophila* 12 Genomes Consortium 2007; Garrigan et al. 2012, 2014; Rogers et al. 2014; Turissini et al. 2015; Melvin et al. 2018; Miller et al. 2018; Schrider et al. 2018; Cooper et al. 2019; Kang et al. 2019; Lanno et al. 2019; Meany et al. 2019; Stewart and Rogers 2019). For an overview of all used publicly available

data, see supplementary table 6, Supplementary Material online.

DNA Sequencing
DNA for Illumina paired-end sequencing was extracted from whole bodies of 20–30 virgin female flies using a salt-extraction protocol (Maniatis et al. 1982). Libraries were prepared with the NEBNext Ultra II DNA library Prep Kit (New England Bioloabs, Ipswich, MA) using 1 μg DNA. Illumina sequencing was performed by the Vienna Biocenter Core Facilities using the HiSeq2500 platform (2 × 125 bp; Illumina, San Diego, CA).

Small RNA Sequencing
For small RNA sequencing, we extracted total RNA from ovaries of the strains Canton-S, Iso-1, and Lausanne-S using TRIzol. The small RNA was sequenced by Fasteris (Geneva, Switzerland). After depletion of 2S rRNA, library preparation was performed using the Illumina TruSeq small RNA kit and cDNA was sequenced on an Illumina NextSeq platform (50 bp; Illumina, San Diego, CA). Adapter sequences were trimmed with cutadapt (v2.3) (Martin 2011) (adapter: TGGAATTCTCGGGTGCCAAGGAACTCCAGTCACCATTTT ATCTCGTATGC) and filtered for reads with a length between 18 and 36 nt. The reads were mapped to a database consisting of *D. melanogaster* miRNAs, mRNAs, rRNAs, snRNAs, snoRNAs, tRNAs (Thurmond et al. 2019), and the TE sequences (Quesneville et al. 2005) using novoalign (v3.09; http://novocraft.com/, last accesssed December 15, 2020) and allowing for two mismatches (unless mentioned otherwise). Solely piRNAs with a length between 23 and 29 nt were retained and the abundance of piRNAs was normalized to a million miRNAs as described previously (Kofler et al. 2018). For computing the ping-pong signatures and visualizing the piRNA abundance along the Tirant sequence, we used a previously developed pipeline (Kofler et al. 2018). To calculate the abundance of piRNAs complementary to the degraded Tirant fragments, we first extracted the sequences of degraded Tirant insertions (>10% divergence to consensus sequence) from the reference assembly of Iso-1 (v6.22) with RepeatMasker (open-4.0.7; Smit et al. 2013–2015) and bedtools (Quinlan and Hall 2010) (v2.29.2). All sequences longer than 100 bp were concatenated (the reverse complement was adjusted with bedtools) and small RNAs were mapped to these sequences using novoalign. The abundance of all piRNAs complementary to degraded Tirant sequences was summed. We also analyzed the small RNA content of the five GDL strains B10, I06, N10, T05, and ZW155 (data are publicly available; Luo et al. 2020).

TE Abundance and Diversity
The coverage along a TE and the frequencies of SNPs and indels in a TE were computed using our newly developed tool DeviaTE (v0.3.8) (Weilguny and Kofler 2019). Briefly, short reads from a sample were aligned with bwa sw (v0.7.17) (Li and Durbin 2009) to the TE consensus sequences of *Drosophila* (Quesneville et al. 2005) as well as to three single-copy genes (*traffic jam*, *rpl32*, and *rhino*), which allowed

us to infer TE copy numbers by contrasting the coverage of a TE to the coverage of the single-copy genes. The abundance and diversity of TE insertions were visualized with DeviaTE (Weilguny and Kofler 2019). To obtain the normalized number of reads mapping to each TE (rpm), we used PopoolationTE2 (v1.10.03) (Kofler et al. 2016). Based on the visualization of the TE composition with DeviaTE and the estimates of the TE abundance (rpm and DeviaTE using normalization with single-copy genes), we manually classified the presence/absence of Tirant, hobo, the I-element, and the P-element in different *D. melanogaster* strains (supplementary table 1, Supplementary Material online). We used the following three categories: 1) absence of any TE sequences, 2) solely degraded TE sequences are present, 3) nondegraded sequences, with a high similarity to the consensus sequence, are present. For example, see supplementary figures 4 and 6–8, Supplementary Material online. A PCA based on the allele frequencies of SNPs in a TE supports our classification for Tirant and hobo. Since many strains do not contain any P-element sequences, the allele frequencies of SNPs in the P-element could not be calculated for all strains. Despite discernible differences between strains with and without recent I-element insertions, the PCA did not separate these two groups (supplementary figs. 6 and 14, Supplementary Material online). The PCA was performed in R (prcomp) using arcsine and square root transformed allele frequencies of SNPs in TEs (R Core Team 2012). The DSPR lines were not included into the PCA due to their short-read length (50 bp). The pairwise F_{ST} based on the SNPs of TEs was computed with Popoolation2 (v1.2.01) (Kofler et al. 2011) ("fst-sliding.pl" –window-size 8526 –max-coverage 0.1%).

We used PoMo (Schrempf et al. 2016) based on the allele frequencies of Tirant SNPs to generate a tree of the species in the *D. melanogaster* species subgroup. PoMo uses allele frequency data to account for the intraspecific differences while calculating the interspecific variation. We run PoMo with IQ-TREE (v1.6.12) (Nguyen et al. 2015) using polymorphism-aware models (HKY + P). We obtained bootstrap estimates for each node using the ultra-fast bootstrap (-bb) option for 1000 replicates.

Tirant sequences in the assemblies of Canton-S (Wierzbicki et al. 2020) and Iso-1 (v6.22; https://flybase.org/, last accessed December 15, 2020) were identified with RepeatMasker using the TE consensus sequences of *Drosophila* as custom library (Quesneville et al. 2005). To visualize the divergence of annotated Tirant fragments of the Canton-S genome, we extract all sequences annotated with RepeatMasker and map them to the Tirant consensus sequence using bwa sw (Li and Durbin 2009) with a low mismatch penalty (-b) of 0.5. Visualization of the sequence alignment was done with IGV. Colored lines represent SNPs compared with the consensus sequence.

We searched for canonical Tirant insertions in a long-read based assembly of *D. simulans* (strain w^{XD1}; PRJNA383250; Chakraborty et al. 2020) using RepeatMasker (open-4.0.7; Smit et al. 2013–2015). We filtered for complete insertions with a low divergence (<5%).

To estimate the position and population frequency of canonical and degraded Tirant insertions in a natural *D. melanogaster* population, we used PoPoolationTE2 (v1.10.03) (Kofler et al. 2016) and a population collected in 2014 at Viltain (France) by the DrosEU consortium (SRR5647729; Kapun et al. 2020). We generated the artificial reference genome required by PoPoolationTE2, by merging the repeat masked reference genome, the consensus sequence of Tirant and the degraded Tirant sequences with a minimum length of 100 bp (see above) into a single fasta file. The short reads were mapped to this artificial reference genome using bwa mem (v0.7.17) (Li and Durbin 2009) with paired-end mode and the parameter -M. The mapped reads were sorted with samtools (Li, Handsaker, et al. 2009). Finally, we followed the PoPoolationTE2 pipeline using the parameters: –map-qual 15, –min-count 2, –min-coverage 2. We indicated heterochromatic regions following previous work (Riddle et al. 2011; Hoskins et al. 2015).

Hybrid Dysgenesis Assay

To test whether Tirant induces HD symptoms, we performed four reciprocal crosses among *D. melanogaster* strains having canonical Tirant insertions (Urbana-S, Hikone-R, and Iso-1) and strains not having canonical Tirant insertions (Lausanne-S, Canton-S, Crimea). Each cross was performed in three replicates by mating 20 female virgin flies with 15 males. To estimate the number of dysgenic ovaries, 2–5 days old F1 flies (kept at either 20, 25, or 29 °C) were allowed to lay eggs on black agar plates (containing charcoal) for 24 h. The F1 female ovaries were dissected on PBS and scored for the presence of dysgenic (underdeveloped) ovaries. The deposited F2 embryos were counted, incubated for 24 h, and the number of larvae (=hatched eggs) was quantified. Crosses involving Iso-1 were only performed at 25 °C.

Author Contributions

F.S. and R.K. conceived the work. F.S. and F.W. analyzed the data. K.-A.S. provided feedback on the manuscript. F.S. and R.K. wrote the manuscript.

Data Availability

All scripts are available at https://sourceforge.net/projects/te-tools/ (last accesssed December 15, 2020) (folder tirant) and important files (including all DeviaTE outputs) at https://sourceforge.net/projects/te-tools/files/tirant_data/ (last

accesssed December 15, 2020). The sequence data of the old laboratory strains and the piRNA sequences are available at NCBI (PRJNA634847).

References

Agis M, Schlötterer C. 2001. Microsatellite variation in natural *Drosophila melanogaster* populations from New South Wales (Australia) and Tasmania. *Mol Ecol.* 10(5):1197–1205.

Akkouche A, Grentzinger T, Fablet M, Armenise C, Burlet N, Braman V, Chambeyron S, Vieira C. 2013. Maternally deposited germline piRNAs silence the tirant retrotransposon in somatic cells. *EMBO Rep.* 14(5):458–464.

Akkouche A, Rebollo R, Burlet N, Esnault C, Martinez S, Viginier B, Terzian C, Vieira C, Fablet M. 2012. tirant, a newly discovered active endogenous retrovirus in *Drosophila simulans*. *J Virol.* 86(7):3675–3681.

Anxolabéhère D, Kidwell MG, Periquet G. 1988. Molecular characteristics of diverse populations are consistent with the hypothesis of a recent invasion of *Drosophila melanogaster* by mobile P elements. *Mol Biol Evol.* 5(3):252–269.

Arguello JR, Laurent S, Clark AG, Gaut B. 2019. Demographic history of the human commensal *Drosophila melanogaster*. *Genome Biol Evol.* 11(3):844–854.

Barckmann B, El-Barouk M, Pélisson A, Mugat B, Li B, Franckhauser C, Fiston Lavier A-S, Mirouze M, Fablet M, Chambeyron S. 2018. The somatic piRNA pathway controls germline transposition over generations. *Nucleic Acids Res.* 46(18):9524–9536.

Bargues N, Lerat E. 2017. Evolutionary history of LTR-retrotransposons among 20 *Drosophila* species. *Mob DNA.* 8(1):1–15.

Bartolomé C, Bello X, Maside X. 2009. Widespread evidence for horizontal transfer of transposable elements across *Drosophila* genomes. *Genome Biol.* 10(2):R22.

Bartolomé C, Maside X, Charlesworth B. 2002. On the abundance and distribution of transposable elements in the genome of *Drosophila melanogaster*. *Mol Biol Evol.* 19(6):926–937.

Begun DJ, Aquadro CF. 1995. Molecular variation at the vermilion locus in geographically diverse populations of *Drosophila melanogaster* and *D. simulans*. *Genetics* 140(3):1019–1032.

Bergland AO, Behrman EL, O'Brien KR, Schmidt PS, Petrov DA. 2014. Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in *Drosophila*. *PLoS Genet.* 10(11):e1004775.

Bergland AO, Tobler R, González J, Schmidt P, Petrov D. 2016. Secondary contact and local adaptation contribute to genome-wide patterns of clinal variation in *Drosophila melanogaster*. *Mol Ecol.* 25(5):1157–1174.

Bergman CM, Bensasson D. 2007. Recent LTR retrotransposon insertion contrasts with waves of non-LTR insertion since speciation in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A.* 104(27):11340–11345.

Biémont C. 2010. A brief history of the status of transposable elements: from junk DNA to major players in evolution. *Genetics* 186(4):1085–1093.

Bingham PM, Kidwell MG, Rubin GM. 1982. The molecular basis of P-M hybrid dysgenesis: the role of the P element, a P-strain-specific transposon family. *Cell* 29(3):995–1004.

Black DM, Jackson MS, Kidwell MG, Dover GA. 1987. KP elements repress P-induced hybrid dysgenesis in *Drosophila melanogaster*. *EMBO J.* 6(13):4125–4135.

Blackman RK, Grimaila R, Macy M, Koehler D, Gelbart WM. 1987. Mobilization of hobo elements residing within the decapentaplegic gene complex: suggestion of a new hybrid dysgenesis system in *Drosophila melanogaster*. *Cell* 49(4):497–505.

Blumenstiel JP. 2011. Evolutionary dynamics of transposable elements in a small RNA world. *Trends Genet.* 27(1):23–31.

Blumenstiel JP. 2019. Birth, school, work, death and resurrection: the life stages and dynamics of transposable element proliferation. *Genes* 10(5):336.

Bock I, Parsons P. 1981. Species of Australia and New Zealand. In Ashburner M, Carson L, Thompson JJ, editors. Vol. 3a. The genetics and biology of Drosophila. Oxford: Academic Press. p. 349–393.

Bonnivard E, Bazin C, Denis B, Higuet D. 2000. A scenario for the hobo transposable element invasion, deduced from the structure of natural populations of *Drosophila melanogaster* using tandem TPE repeats. *Genet Res.* 75(1):13–23.

Bowen NJ, McDonald JF. 2001. *Drosophila* euchromatic LTR retrotransposons are much younger than the host species in which they reside. *Genome Res.* 11(9):1527–1540.

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 128(6):1089–1103.

Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ. 2008. An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* 322(5906):1387–1392.

Brizuela BJ, Elfring L, Ballard J, Tamkun JW, Kennison JA. 1994. Genetic analysis of the *brahma* gene of *Drosophila melanogaster* and polytene chromosome subdivisions 72AB. *Genetics* 137(3):803–813.

Bucheton A. 1979. Non-Mendelian female sterility in *Drosophila melanogaster*: influence of aging and thermic treatments. III. Cumulative effects induced by these factors. *Genetics* 93(1):131–142.

Bucheton A, Lavige JM, Picard G, L'Heritier P. 1976. Non-Mendelian female sterility in *Drosophila melanogaster*: quantitative variations in the efficiency of inducer and reactive strains. *Heredity* 36(3):305–314.

Bucheton A, Paro R, Sang HM, Pelisson A, Finnegan DJ. 1984. The molecular basis of I-R hybrid Dysgenesis in *Drosophila melanogaster*: identification, cloning, and properties of the I factor. *Cell* 38(1):153–163.

Bucheton A, Simonelig M, Vaury C, Crozatier M. 1986. Sequences similar to the I transposable element involved in I-R hybrid dysgenesis in *D. melanogaster* occur in other *Drosophila* species. *Nature* 322(6080):650–652.

Bucheton A, Vaury C, Chaboissier M-C, Abad P, Pélisson A, Simonelig M. 1992. I elements and the *Drosophila* genome. *Genetica* 86(1-3):175–190.

Calvi BR, Gelbart WM. 1994. The basis for germline specificity of the hobo transposable element in *Drosophila melanogaster*. *EMBO J.* 13(7):1636–1644.

Cañizares J, Grau M, Paricio N, Moltó MD. 2000. Tirant is a new member of the Gypsy family of retrotransposons in *Drosophila melanogaster*. *Genome* 43(1):9–14.

Chakraborty M, Chang C-H, Khost D, Vedanayagam J, Adrion JR, Liao Y, Montooth KL, Meiklejohn CD, Larracuente AM, Emerson JJ. 2020. Evolution of genome structure in the Drosophila simulans species complex. *bioRxiv*.

Chakraborty M, Emerson JJ, Macdonald SJ, Long AD. 2019. Structural variants exhibit widespread allelic heterogeneity and shape variation in complex traits. *Nat Commun.* 10(1):419275.

Cooper JC, Guo P, Bladen J, Phadnis N. 2019. A triple-hybrid cross reveals a new hybrid incompatibility locus between D. melanogaster and D. sechellia. *bioRxiv* 590588.

Czech B, Preall JB, McGinn J, Hannon GJ. 2013. A transcriptome-wide RNAi screen in the *Drosophila* ovary reveals factors of the germline piRNA pathway. *Mol Cell.* 50(5):749–761.

Daniels SB, Chovnick A, Boussy IA. 1990. Distribution of hobo transposable elements in the genus *Drosophila*. *Mol Biol Evol.* 7(6):589–606.

Daniels SB, Peterson KR, Strausbaugh LD, Kidwell MG, Chovnick A. 1990. Evidence for horizontal transmission of the P transposable element between *Drosophila* species. *Genetics* 124(2):339–355.

Doolittle WF, Sapienza C. 1980. Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284(5757):601–603.

Drosophila 12 Genomes Consortium. 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450(7167):203–218.

Eanes WF, Wesley C, Charlesworth B. 1992. Accumulation of P elements in minority inversions in natural populations of *Drosophila melanogaster*. *Genet Res.* 59(1):1–9.

Engels WR. 1992. The origin of P elements in *Drosophila melanogaster*. *Bioessays* 14(10):681–686.

Engels WR. 2007. Hybrid dysgenesis in *Drosophila melanogaster*: rules of inheritance of female sterility. *Genet Res*. 89(5-6):407–424.

Fablet M, Lerat E, Rebollo R, Horard B, Burlet N, Martinez S, Brasset È, Gilson E, Vaury C, Vieira C. 2009. Genomic environment influences the dynamics of the tirant LTR retrotransposon in *Drosophila*. *FASEB J*. 23(5):1482–1489.

Fablet M, McDonald JF, Biémont C, Vieira C. 2006. Ongoing loss of the tirant transposable element in natural populations of *Drosophila simulans*. *Gene* 375(1–2):54–62.

Fablet M, Souames S, Biémont C, Vieira C. 2007. Evolutionary pathways of the tirant LTR retrotransposon in the *Drosophila melanogaster* subgroup of species. *J Mol Evol*. 64(4):438–447.

Galindo MI, Ladevèze V, Lemeunier F, Kalmes R, Periquet G, Pascual L. 1995. Spread of the autonomous transposable element hobo in the genome of *Drosophila melanogaster*. *Mol Biol Evol*. 12(5):723–734.

Garrigan D, Kingan SB, Geneva AJ, Andolfatto P, Clark AG, Thornton KR, Presgraves DC. 2012. Genome sequencing reveals complex speciation in the *Drosophila simulans* clade. *Genome Res*. 22(8):1499–1511.

Garrigan D, Kingan SB, Geneva AJ, Vedanayagam JP, Presgraves DC. 2014. Genome diversity and divergence in *Drosophila mauritiana*: multiple signatures of faster X evolution. *Genome Biol Evol*. 6(9):2444–2458.

Ginzburg LR, Bingham PM, Yoo S. 1984. On the theory of speciation induced by transposable elements. *Genetics* 107(2):331–341.

Goodier JL. 2016. Restricting retrotransposons: a review. *Mob DNA*. 7(1):16.

Grenier JK, Roman Arguello J, Moreira MC, Gottipati S, Mohammed J, Hackett SR, Boughton R, Greenberg AJ, Clark AG. 2015. Global diversity lines-a five-continent reference panel of sequenced *Drosophila melanogaster* strains. *G3 (Bethesda)* 5(4):593–603.

Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC. 2007. A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in *Drosophila*. *Science* 315(5818):1587–1590.

Hickey DA. 1982. Selfish DNA: a sexually transmitted nuclear parasite. *Genetics* 101(3-4):519–531.

Hoskins RA, Carlson JW, Wan KH, Park S, Mendez I, Galle SE, Booth BW, Pfeiffer BD, George RA, Svirskas R, et al. 2015. The Release 6 reference sequence of the *Drosophila melanogaster* genome. *Genome Res*. 25(3):445–458.

Jakšić AM, Kofler R, Schlötterer C. 2017. Regulation of transposable elements: interplay between TE-encoded regulatory sequences and host-specific trans-acting factors in *Drosophila melanogaster*. *Mol Ecol*. 26(19):5149–5159.

Kang L, Rashkovetsky E, Michalak K, Garner HR, Mahaney JE, Rzigalinski BA, Korol A, Nevo E, Michalak P. 2019. Genomic divergence and adaptive convergence in *Drosophila simulans* from evolution Canyon, Israel. *Proc Natl Acad Sci U S A*. 116(24):11839–11844.

Kapun M, Barrón MG, Staubach F, Obbard DJ, Wiberg RAW, Vieira J, Goubert C, Rota-Stabelli O, Kankare M, Bogaerts-Márquez M, et al. 2020. Genomic analysis of European *Drosophila melanogaster* populations reveals longitudinal structure, continent-wide selection, and previously unknown DNA viruses. *Mol Biol Evol*. 37(9):2661–2678.

Kelleher E, Jaweria J, Akoma U, Ortega L, Tang W. 2018. QTL mapping of natural variation reveals that the developmental regulator Bruno reduces tolerance to P-element transposition in the Drosophila female germline. *PLoS Biol*. 16(10):e2006040.

Keller A. 2007. *Drosophila melanogaster*'s history as a human commensal. *Curr Biol*. 17(3):R77–R81.

Kidwell MG. 1983. Evolution of hybrid dysgenesis determinants in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A*. 80(6):1655–1659.

Kidwell MG, Kidwell JF, Sved JA. 1977. Hybrid dysgenesis in *Drosophila melanogaster*: a syndrome of aberrant traits including mutations, sterility and male recombination. *Genetics* 86(4):813–833.

Kidwell MG, Kimura K, Black DM. 1988. Evolution of hybrid dysgenesis potential following P element contamination in *Drosophila melanogaster*. *Genetics* 119(4):815–828.

Kidwell MG, Novy JB. 1979. Hybrid dysgenesis in Drosophila melanogaster: sterility resulting from gonadal dysgenesis in the P-M system. *Genetics* 92(4):1127–1140.

King EG, Merkes CM, McNeil CL, Hoofer SR, Sen S, Broman KW, Long AD, Macdonald SJ. 2012. Genetic dissection of a model complex trait using the *Drosophila* synthetic population resource. *Genome Res*. 22(8):1558–1566.

Kofler R, Betancourt AJ, Schlötterer C. 2012. Sequencing of pooled DNA samples (Pool-Seq) uncovers complex dynamics of transposable element insertions in *Drosophila melanogaster*. *PLoS Genet*. 8(1):e1002487.

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.

Kofler R, Hill T, Nolte V, Betancourt AJ, Schlötterer C. 2015. The recent invasion of natural *Drosophila simulans* populations by the P-element. *Proc Natl Acad Sci U S A*. 112(21):6659–6663.

Kofler R, Nolte V, Schlötterer C. 2015. Tempo and mode of transposable element activity in *Drosophila*. *PLoS Genet*. 11(7):e1005406–e1005421.

Kofler R, Pandey RV, Schlötterer C. 2011. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* 27(24):3435–3436.

Kofler R, Senti K-A, Nolte V, Tobler R, Schlotterer C. 2018. Molecular dissection of a natural transposable element invasion. *Genome Res*. 28(6):824–835.

Kotov AA, Adashev VE, Godneeva BK, Ninova M, Shatskikh AS, Bazylev SS, Aravin AA, Olenina LV. 2019. PiRNA silencing contributes to interspecies hybrid sterility and reproductive isolation in *Drosophila melanogaster*. *Nucleic Acids Res*. 47(8):4255–4271.

Lack JB, Cardeno CM, Crepeau MW, Taylor W, Corbett-Detig RB, Stevens KA, Langley CH, Pool JE. 2015. The *Drosophila* genome nexus: a population genomic resource of 623 *Drosophila melanogaster* genomes, including 197 from a single ancestral range population. *Genetics* 199(4):1229–1241.

Lanno SM, Shimshak SJ, Peyser RD, Linde SC, Coolon JD. 2019. Investigating the role of Osiris genes in *Drosophila sechellia* larval resistance to a host plant toxin. *Ecol Evol*. 9(4):1922–1933.

Le Rouzic A, Capy P. 2005. The first steps of transposable elements invasion: parasitic strategy vs. genetic drift. *Genetics* 169(2):1033–1043.

Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Tóth KF. 2013. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev*. 27(4):390–399.

Le Thomas A, Stuwe E, Li S, Du J, Marinov G, Rozhkov N, Chen YCA, Luo Y, Sachidanandam R, Toth KF, et al. 2014. Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. *Genes Dev*. 28(15):1667–1680.

Lemeunier F, Ashburner M, Thoday JM. 1976. Relationships within the *melanogaster* species subgroup of the genus *Drosophila* (*Sophophora*) – II. Phylogenetic relationships between six species based upon polytene chromosome banding sequences. *Proc R Soc Lond B Biol Sci*. 193(1112):275–294.

Lerat E, Burlet N, Biémont C, Vieira C. 2011. Comparative analysis of transposable elements in the *melanogaster* subgroup sequenced genomes. *Gene* 473(2):100–109.

Li C, Vagin VV, Lee S, Xu J, Ma S, Xi H, Seitz H, Horwich MD, Syrzycka M, Honda BM, et al. 2009. Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell* 137(3):509–521.

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

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16):2078–2079.

Lindsley DH, Grell EH. 1968. Genetic variations of Drosophila melanogaster. Washington: Carnegie Institute of Washington Publication.

Lohe AR, Moriyama EN, Lidholm DA, Hartl DL. 1995. Horizontal transmission, vertical inactivation, and stochastic loss of mariner-like transposable elements. *Mol Biol Evol*. 12(1):62–72.

Loreto EL, Carareto CMA, Capy P. 2008. Revisiting horizontal transfer of transposable elements in *Drosophila*. *Heredity* 100(6):545–554.

Luo S, Zhang H, Duan Y, Yao X, Clark AG, Lu J. 2020. The evolutionary arms race between transposable elements and piRNAs in *Drosophila melanogaster*. *BMC Evol Biol*. 20(1):1–18.

Machado HE, Bergland AO, Taylor R, Tilk S, Behrman E, Dyer K, Fabian DK, Flatt T, González J, Karasov TL. 2019. Broad geographic sampling reveals predictable, pervasive, and strong seasonal adaptation in Drosophila. *bioRxiv*, 337543.

Mackay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, Casillas S, Han Y, Magwire MM, Cridland JM, et al. 2012. The *Drosophila melanogaster* genetic reference panel. *Nature* 482(7384):173–178.

Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ. 2009. Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* 137(3):522–535.

Maniatis T, Fritsch EF, Sambrook J, et al. 1982. Molecular cloning: a laboratory manual. Vol. 545. NY: Cold Spring Harbor Laboratory.

Marí-Ordóñez A, Marchais A, Etcheverry M, Martin A, Colot V, Voinnet O. 2013. Reconstructing *de novo* silencing of an active plant retrotransposon. *Nat Genet*. 45(9):1029–1039.

Marsano RM, Moschetti R, Caggese C, Lanave C, Barsanti P, Caizzi R. 2000. The complete Tirant transposable element in *Drososphila melanogaster* shows a structural relationship with retrovirus-like retrotransposons. *Gene* 247(1–2):87–95.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 17(1):10–12.

Maruyama K, Hartl DL. 1991. Evidence for interspecific transfer of the transposable element mariner between *Drosophila* and *Zaprionus*. *J Mol Evol*. 33(6):514–524.

Meany MK, Conner WR, Richter SV, Bailey JA, Turelli M, Cooper BS. 2019. Loss of cytoplasmic incompatibility and minimal fecundity effects explain relatively low Wolbachia frequencies in *Drosophila mauritiana*. *Evolution* 73(6):1278–1295.

Melvin RG, Lamichane N, Havula E, Kokki K, Soeder C, Jones CD, Hietakangas V. 2018. Natural variation in sugar tolerance associates with changes in signaling and mitochondrial ribosome biogenesis. *eLife* 7:e40841.

Miller DE, Staber C, Zeitlinger J, Hawley RS. 2018. Highly contiguous genome assemblies of 15 *Drosophila* species generated using nanopore sequencing. *G3 (Bethesda)* 8(10):3131–3141.

Mizrokhi LJ, Mazo AM. 1990. Evidence for horizontal transmission of the mobile element jockey between distant *Drosophila* species. *Proc Natl Acad Sci U S A*. 87(23):9216–9220.

Moltó MD, Paricio N, López-Preciado M. A, Semeshin VF, Martínez-Sebastián MJ. 1996. Tirant: a new retrotransposon-like element in *Drosophila melanogaster*. *J Mol Evol*. 42(3):369–375.

Moon S, Cassani M, Lin YA, Wang L, Dou K, Zhang ZZ. 2018. A robust transposon-endogenizing response from germline stem cells. *Dev Cell*. 47(5):660–671.e3.

Mugnier N, Gueguen L, Vieira C, Biémont C. 2008. The heterochromatic copies of the LTR retrotransposons as a record of the genomic events that have shaped the *Drosophila melanogaster* genome. *Gene* 411(1–2):87–93.

Nefedova LN, Urusov FA, Romanova NI, Shmel'Kova AO, Kim AI. 2012. Study of the transcriptional and transpositional activities of the Tirant retrotransposon in *Drosophila melanogaster* strains mutant for the flamenco locus. *Genetika* 48(11):1271–1279.

Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 32(1):268–274.

Orgel LE, Crick FH. 1980. Selfish DNA: the ultimate parasite. *Nature* 284(5757):604–607.

Pardue ML, DeBaryshe PG. 2011. Retrotransposons that maintain chromosome ends. *Proc Natl Acad Sci U S A*. 108(51):20317–20324.

Parhad SS, Tu S, Weng Z, Theurkauf WE. 2017. Adaptive evolution leads to cross-species incompatibility in the piRNA transposon silencing machinery. *Dev Cell*. 43(1):60–70.e5.

Parsons P, Stanley S. 1981. Special ecological studies-domesticated and widespread species. In: Ashburner M, Carson L, Thompson JJ, editors. Vol. 3c. The genetics and biology of Drosophila. Oxford: Academic Press. p. 349–393.

Pascual L, Periquet G. 1991. Distribution of hobo transposable elements in natural populations of *Drosophila melanogaster*. *Mol Biol Evol*. 8(3):282–296.

Peccoud J, Loiseau V, Cordaux R, Gilbert C. 2017. Massive horizontal transfer of transposable elements in insects. *Proc Natl Acad Sci U S A*. 114(18):4721–4726.

Periquet G, Hamelin MH, Bigot Y, Lepissier A. 1989. Geographical and historical patterns of distribution of hobo elements in *Drosophila melanogaster* populations. *J Evol Biol*. 2(3):223–229.

Periquet G, Lemeunier F, Bigot Y, Hamelin MH, Bazin C, Ladevéze V, Eeken J, Galindo MI, Pascual L, Boussy I. 1994. The evolutionary genetics of the hobo transposable element in the *Drosophila melanogaster* complex. *Genetica* 93(1–3):79–90.

Petrov DA, Fiston-Lavier A-S, Lipatov M, Lenkov K, González J. 2011. Population genomics of transposable elements in *Drosophila melanogaster*. *Mol Biol Evol*. 28(5):1633–1644.

Post C, Clark JP, Sytnikova YA, Chirn G-W, Lau NC. 2014. The capacity of target silencing by *Drosophila* PIWI and piRNAs The capacity of target silencing by *Drosophila* PIWI and piRNAs. *RNA* 20(12):1977–1986.

Quesneville H, Bergman CM, Andrieu O, Autard D, Nouaud D, Ashburner M, Anxolabéhère D. 2005. Combined evidence annotation of transposable elements in genome sequences. *PLoS Comp Biol*. 1(2):e22–e175.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6):841–842.

R Core Team. 2012. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing.

Rahman R, Chirn GW, Kanodia A, Sytnikova YA, Brembs B, Bergman CM, Lau NC. 2015. Unique transposon landscapes are pervasive across *Drosophila melanogaster* genomes. *Nucleic Acids Res*. 43(22):10655–10672.

Riddle NC, Minoda A, Kharchenko PV, Alekseyenko AA, Schwartz YB, Tolstorukov MY, Gorchakov AA, Jaffe JD, Kennedy C, Linder-Basso D, et al. 2011. Plasticity in patterns of histone modifications and chromosomal proteins in *Drosophila* heterochromatin. *Genome Res*. 21(2):147–163.

Rogers RL, Cridland JM, Shao L, Hu TT, Andolfatto P, Thornton KR. 2014. Landscape of standing variation for tandem duplications in *Drosophila yakuba* and *Drosophila simulans*. *Mol Biol Evol*. 31(7):1750–1766.

Rozhkov NV, Hammell M, Hannon GJ. 2013. Multiple roles for Piwi in silencing *Drosophila* transposons. *Genes Dev*. 27(4):400–412.

Ruebenbauer A, Schlyter F, Hansson BS, Löfstedt C, Larsson MC. 2008. Genetic variability and robustness of host odor preference in *Drosophila melanogaster*. *Curr Biol*. 18(18):1438–1443.

Saint-Leandre B, Levine MT. 2020. The telomere paradox: stable genome preservation with rapidly evolving proteins. *Trends Genet*. 36(4):232–242.

Sánchez-Gracia A, Maside X, Charlesworth B. 2005. High rate of horizontal transfer of transposable elements in *Drosophila*. *Trends Genet*. 21(4):200–203.

Schrempf D, Minh BQ, De Maio N, von Haeseler A, Kosiol C. 2016. Reversible polymorphism-aware phylogenetic models and their application to tree inference. *J Theor Biol*. 407:362–370.

Schrider D, Ayroles J, Matute D, Kern A. 2018. Supervised machine learning reveals introgressed loci in the genomes of *Drosophila simulans* and *D. sechellia*. *PLoS Genet*. 14(4):e1007341.

Serrato-Capuchina A, Wang J, Earley E, Peede D, Isbell K, Matute D. 2020. Paternally inherited P-element copy number affects the magnitude of hybrid dysgenesis in *Drosophila simulans* and *D. melanogaster*. *Genome Biol Evol*. 12(6):808–826.

Sienski G, Dönertas D, Brennecke J. 2012. Transcriptional silencing of transposons by Piwi and Maelstrom and its impact on chromatin state and gene expression. *Cell* 151(5):964–980.

Silva JC, Loreto EL, Clark JB. 2004. Factors that affect the horizontal transfer of transposable elements. *Curr Issues Mol Biol.* 6(1):57–72.

Simmons GM. 1992. Horizontal transfer of hobo transposable elements within the *Drosophila melanogaster* species complex: evidence from DNA sequencing. *Mol Biol Evol.* 9(6):1050–1060.

Smit AFA, Hubley R, Green P. 2013–2015. RepeatMasker Open-4.0.

Sniegowski PD, Charlesworth B. 1994. Transposable element numbers in cosmopolitan inversions from a natural population of *Drosophila melanogaster*. *Genetics* 137(3):815–827.

Song SU, Kurkulos M, Boeke JD, Corces VG. 1997. Infection of the germ line by retroviral particles produced in the follicle cells: a possible mechanism for the mobilization of the Gypsy retroelement of *Drosophila*. *Development* 124(14):2789–2798.

Srivastav SP, Kelleher ES. 2017. Paternal induction of hybrid dysgenesis in *Drosophila melanogaster* is weakly correlated with both P-element and Hob element dosage. *G3 (Bethesda)* 7(5):1487–1497.

Srivastav SP, Rahman R, Ma Q, Pierre J, Bandyopadhyay S, Lau NC. 2019. Har-P, a short P-element variant, weaponizes P-transposase to severely impair *Drosophila* development. *eLife* 26(6):e49948.

Stewart NB, Rogers RL. 2019. Chromosomal rearrangements as a source of new gene formation in *Drosophila yakuba*. *PLoS Genet.* 15(9):e1008314.

Terzian C, Ferraz C, Demaille J, Bucheton A. 2000. Evolution of the Gypsy endogenous retrovirus in the *Drosophila melanogaster* subgroup. *Mol Biol Evol.* 17(6):908–914.

Terzian C, Pélisson A, Bucheton A. 2001. Evolution and phylogeny of insect endogenous retroviruses. *BMC Evol Biol.* 1(1):3.

Thurmond J, Goodman JL, Strelets VB, Attrill H, Gramates LS, Marygold SJ, Matthews BB, Millburn G, Antonazzo G, Trovisco V, et al. 2019. FlyBase 2.0: the next generation. *Nucleic Acids Res.* 47(D1):D759–D765.

Turissini DA, Liu G, David JR, Matute DR. 2015. The evolution of reproductive isolation in the *Drosophila yakuba* complex of species. *J Evol Biol.* 28(3):557–575.

Vieira C, Lepetit D, Dumont S, Biémont C. 1999. Wake up of transposable elements following *Drosophila simulans* worldwide colonization. *Mol Biol Evol.* 16(9):1251–1255.

Viggiano L, Caggese C, Barsanti P, Caizzi R. 1997. Cloning and characterization of a copy of Tirant transposable element in *Drosophila melanogaster*. *Gene* 197(1-2):29–35.

Wang L, Barbash DA, Kelleher ES. 2020. Adaptive evolution among cytoplasmic piRNA proteins leads to decreased genomic autoimmunity. *PLoS Genet.* 16(6):e1008861–e1008922.

Wang L, Dou K, Moon S, Tan FJ, Zhang ZZZ, Wang L, Dou K, Moon S, Tan FJ, Zhang ZZZ. 2018. Hijacking oogenesis enables massive propagation article hijacking oogenesis enables massive propagation of LINE and retroviral transposons. *Cell* 174(5):1082–1094.

Weilguny L, Kofler R. 2019. DeviaTE: assembly-free analysis and visualization of mobile genetic element composition. *Mol Ecol Resour.* 19(5):1346–1354.

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, et al. 2007. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet.* 8(12):973–982.

Wierzbicki F, Schwarz F, Cannalonga O, Kofler R. 2020. Generating high quality assemblies for genomic analysis of transposable elements. *bioRxiv*.

Yang P, Wang Y, Macfarlan TS. 2017. The role of KRAB-ZFPs in transposable element repression and mammalian evolution. *Trends Genet.* 33(11):871–881.

Yannopoulos G, Stamatis N, Monastirioti M, Hatzopoulos P, Louis C. 1987. hobo is responsible for the induction of hybrid dysgenesis by strains of *Drosophila melanogaster* bearing the male recombination factor 23.5MRF. *Cell* 49(4):487–495.