Hybrid dysgenesis in *Drosophila simulans* due to a rapid global invasion of the *P*-element

Tom Hill[1], Christian Schlötterer[1], Andrea J. Betancourt[1]*

[1]Institut für Populationsgenetik, Vetmeduni Vienna, Veterinärplatz 1, A-1210 Wien, Austria.

Phone: +43 1 250 77 4334

Fax: +43 1 250 77 4390

Email: andrea.betancourt@vetmeduni.ac.at

Keywords: *P*-element, *Drosophila*, hybrid dysgenesis, transposable elements

Running Title: Dysgenesis in *simulans* due to the spread of the *P*-element

Abstract

In a classic example of the invasion of a species by a selfish genetic element, the *P*-element was horizontally transferred from a distantly related species into *Drosophila melanogaster*, where it caused a syndrome of abnormal phenotypes, including sterility, called 'hybrid dysgenesis'. The P-element spread globally in the course of a few decades in *D. melanogaster*, while its sister species, including *D. simulans*, remained *P*-element free. Here, we find hybrid dysgenesis also occurs between *D. simulans* strains collected in different years; a survey of 154 strains shows that over one-third of them induce hybrid dysgenesis. Using genomic and transcriptomic data, we show that this dysgenesis-inducing phenotype is associated with the presence of the P-element, which has recently invaded *D. simulans*. We survey 573 *D. simulans* strains collected over the past 30 years for the presence of the *P*-element, and find that the *D. simulans* invasion of the *P*-element occurred rapidly, with infected strains across three continents being rare in 2004 and common by 2014. Importantly, strains collected from the latter phase of this invasion have adapted to ameliorate the effects of the P-element, as evidenced by their resistance to the hybrid dysgenesis phenotype. This work demonstrates that selfish elements can cause rapid global change in the genomes of their host species, even faster than previously thought.

Introduction

Discovered in the 1950's by Barabara McClintock (1), transposable elements (TEs) have been described as the 'ultimate parasite' (2,3). Genetically simple, with autonomous elements consisting of as little as one protein coding gene, but remarkably successful, TEs are responsible for much of the bulk of large eukaryotic genomes (4). They persist by spreading within the genome of a species, potentially damaging host genes and imposing other fitness costs on their hosts as they spread (5,6). In fact, the mutagenic properties of transposons have been harnessed by molecular geneticists (7).

In the short-term, TEs are vertically transmitted, and the damage they can safely cause the host without extinguishing themselves must therefore be limited. In the long-term, however, TEs are also horizontally transmitted (8–10), and may initially cause substantial damage (11,12). The P-element, a DNA transposon, is the best-studied example. Initially discovered in *Drosophila melanogaster* due to a phenomenon called 'hybrid dysgenesis' (13–16)— a syndrome of phenotypes attributable to DNA damage caused by active P-elements— the P-element had been recently horizontally acquired from a distant relative (17). Once acquired by *D. melanogaster*, the P-element apparently spread rapidly through the species, as evidenced by the fact that it is mostly absent from old laboratory stocks (13,18,19). In its native species, and in *D. melanogaster* collected after the spread of the P-element, its activity is suppressed (17,20). Naïve laboratory strains, on the other hand, were derived from flies collected before the spread of the P-element. As they had not yet evolved the ability to suppress the P-element, their offspring suffer from hybrid dysgenesis under some conditions, rendering them unfit and potentially sterile (12,13).

Since its initial discovery, P-element has become the workhorse of modern Drosophila genetics, providing the basis for genetic manipulation and the previously mentioned mutagenesis (7,21). Until recently, the P-element did not occur in *D. melanogaster's* close relatives, but it has been recently horizontally acquired by *D. simulans* (22,23). Here, we chronicle the entire invasion of this selfish transposon as it spreads through natural populations of *D. simulans*; through its initial deleterious phase, and subsequent suppression in recently collected strains.

Results & Discussion

Hybrid (or gonadal) dysgenesis is a phenomenon in *Drosophila melanogaster* first identified in the F1 offspring of males derived from more recently collected 'P-type' strains and females from old laboratory 'M-type' strains (13,16). This syndrome involves a number of abnormal phenotypes, including spontaneous male recombination, increased mutation rates and dysgenic (abnormally small) gonads (24). Subsequently, Bingham *et al.* (15) discovered that P-type strains harbor a transposable element (TE) absent from the M-types, the *P*-element, and that dysgenesis is due to actively transposing *P*-element in the zygote (15,17). Without any suppression, as in M♀ x P♂ crosses, transposition can result in excessive chromosome breakage and apoptosis of cells in the developing germ-line, frequently yielding sterile offspring (16). The reciprocal crosses, with the P-type as the maternal strain, produce normal offspring (13,24). It was eventually discovered that these offspring are protected by small RNAs (PIWI interacting RNAs or piRNAs) complementary to the *P*-element, which are sequestered in the cytoplasm of the P-type egg, and which silence the *P*-element mRNA (25). In the eggs of M-type strains, these piRNAs are absent, resulting in active transposition of *P*-elements in the gonads

of the dysgenic offspring (Figure 1) (25). Since that time, two other hybrid dysgenic systems have been discovered in *D. melanogaster*, and at least one in *D. virilis* (26–28).

Here, we assayed for hybrid dysgenesis in crosses between strains in a sister species of *D. melanogaster*, *D. simulans*, which did not contain the *P*-element until recently (22,23). Our initial assay used 22 strains collected in Georgia in 2009 [generously donated by P.H. Haddrill, (29)] and 12 strains collected in Madagascar in 2004 & 2008 (collected by B. Ballard and J. David; see Supplementary Table 1 for all fly strains used). We paired strains from the two populations randomly, such that each strain was involved in several crosses (at least 5 for the Madagascar strains and at least 2 for the Georgia strains). Among these crosses, 38.7% showed unidirectional dysgenesis in F1 females at 29°C, a temperature which induces hybrid dysgenesis in *D. melanogaster*, but not at 25°C (36 of 93 crosses showed a significant difference between reciprocals at the 5% level by Fisher's Exact Test; Supplementary Figure 1). Differences between the F1 females from reciprocal crosses, which have identical nuclear genotypes, are probably due to different maternal contributions to the cytoplasm, implicating a cyto-nuclear incompatibility, as in the hybrid dysgenesis phenotype described above (16,24,30). An additional candidate factor is *Wolbachia:* these strains differ in the presence of a *Wolbachia* strain that causes embryonic lethality in one direction of the cross (29,31), and which may potentially cause other differences between reciprocal crosses. However, crosses repeated with untreated strains and the same strains cured of *Wolbachia* infection showed qualitatively similar results (Supplementary Figure 1, Fisher's Exact Test $p > 0.05$ for all comparisons); additional cytotyping crosses described below were performed after tetracycline curing.

Four of the Georgia lines consistently showed dysgenic offspring when used as the paternal line in a cross. Control crosses between these "dysgenesis-inducing" (DI) lines showed no significant unidirectional dysgenesis (Supplementary Figure 1), nor did crosses of these lines to three of the other Georgia lines (Supplementary Figure 1, Fisher's Exact Test $p < 0.05$), implying that 17 of 22 Georgia strains are resistant to the dysgenesis-inducing factor. In contrast, 10 of 10 strains from Madagascar were "dysgenesis-susceptible" (DS) when crossed to DI lines; crosses between these lines also resulted in no hybrid dysgenesis (Supplementary Figure 1). The DI and DS classes are analogous to the P-type and M-type of *D. melanogaster,* respectively (16).

Using DI and DS strains, we cytotyped additional strains [Figure 2; (23,32)]. We reciprocally crossed each strain to at least one known DI and DS strain, and roughly classified tested strains into DI or DS types, based on the phenotypes of the female F1 offspring. We also observed a third type of strain, "dysgenesis-resistant" (DR), which produces no dysgenic offspring in any cross, corresponding to the Q and KP-types in *D. melanogaster* that repress transposition of the *P*-element, but which do not themselves contain active elements (12,25,33,34). Typically, these strains contain degraded copies of the dysgenesis-causing element which produce piRNAs complementary to the TE (25,35).

In all, we cytotyped 154 additional strains from 23 populations sampled between 2002 and 2014 (Figure 2; Supplementary Figure 2). Overall, 21.5% of strains were DI, 25.8% DR, and 52.7% DS. Fertility assays performed on a subset of these crosses show that female offspring of the DI♂ x DS♀ cross (*i.e.,* the dysgenic cross) produced significantly fewer offspring that the reciprocal crosses, consistent with a reduction in their fertility (Supplementary Figure 3; Wilcoxon Rank Sum $p < 0.05$).

DI and DR strains increase in frequency relative to DS strains over time, consistent with the spread of a dysgenesis-causing factor and associated resistance type (Table 1; with DI and DR classes grouped together, both classes explain significant variation in the Americas and Africa, binomial generalized linear model (GLM); Africa: $z = 3.474, p = 0.0051$, Americas: $z = 3.741, p = 0.0002$].

The *D. simulans* hybrid dysgenesis phenotype is due to the *P*-element

Given the similarity between transposable element-induced dysgenesis in *D. melanogaster* and the dysgenic phenotypes seen here, we suspected that the DI lines harbor a dysgenesis-inducing transposable element absent from the DS lines. We therefore examined the DI and DS lines for differences in transposable element content. To identify candidate TEs, we used RNAseq data obtained from a pooled sample of lines from the Florida population [originally collected in 2010, data from (23)], which contains mostly strains with DI and DR cytotypes; we found 68 TEs expressed in this sample.

We then investigated whether any of these elements appear to cause the observed hybrid dysgenesis. To do this, we used genome sequence data from 13 single F1 females produced from crosses of males of strains from the same Florida population used for the RNAseq data and to females from a sequenced reference strain from Madagascar [M252: a DS strain, collected in Madagascar in 2004; (36)]. We examined these sequences for TE insertions present in the Florida lines, but absent from M252. Only one of the expressed TEs had appreciable coverage (at least the average coverage of the genome) that differed between the F1 flies and the M252 line, the *P*-element (Supplementary Table 2, Supplementary Figure 4). With short-read sequence data, we could not unambiguously identify full-length insertions, but

three sequenced strains showed coverage of the complete *P*-element reference sequence for paired-end reads, suggesting that these strains contain full-length *P*-element (23).

Second, we also looked for an effect of *P*-element copy number on the level of dysgenesis, as occurs in *D. melanogaster* (15), and as might be expected if the dysgenic phenotype is induced by the *P*-element. Consistent with this expectation, the fraction of dysgenic F1 females increases with the coverage of the *P*-element in the genome sequence data from crosses where the paternal line is from Florida (with coverage of the P-element standardized by the average coverage of the euchromatic genome; Supplementary figure 5; binomial GLM; $z = 5.87$, $p = 4.25 \times 10^{-9}$), but shows no relationship in the reciprocal cross (binomial GLM; $z = -1.28$, $p = 0.20$; note that we could not distinguish between full-length and partial copies in this analysis). There is no relationship between the dysgenesis phenotype and the copy numbers of other TEs ($p > 0.05$ for all, Supplementary Table 2).

Finally, we designed primers within the exons of the 68 expressed TEs (Supplementary Table 2), and tested for differences between previously cytotyped DI and DS strains using PCR (6 DI strains from Florida 2010, 4 DI strains from Georgia 2009 and 10 DS strains from Madagascar 2004 and 2008). All but two of the elements with appreciable coverage were found to be either present in or absent from all strains (*P*-element and *Tom1*; Supplementary Table 2). *Tom1* shows no significant association with cytotype when additional cytotyped strains were tested (present in 39 of 39 DI strains, and 71 of 78 DS strains, Fisher's Exact Test; $p = 0.19$). The presence of full-length *P*-elements, in contrast, is strongly associated with the DI cytotype: product corresponding to full-length *P*-elements could be amplified from 33 of 39 DI strains using standard PCR, and from 4 of the 6 remaining strains using RT-

PCR. We were able to amplify all 4 exons of the *P*-element individually in all 39 strains. In contrast, no *P*-element could be amplified from any of the 62 DS strains (Fisher's Exact Test on full-length element obtained using standard PCR in DI vs. DS strains; $p = 1.66 \times 10^{-30}$). Using RT-PCR, we confirmed that full-length *P*-element is expressed in the 5 DI strains used in experimental crosses for cytotyping. No full-length transcript could be amplified from the tested DS strains (Supplementary Figure 6). In addition to full-length copies of *P*-element, we find some partial copies, typically missing parts or all of the two internal exons (Supplementary Table 1), consistent with the generation of internal deletions *via* incomplete excision.

Consistent with a homology based silencing mechanism, resistance to hybrid dysgenesis appears to be associated with the presence of the *P*-element in some form, either partial or full-length (we were able to amplify *P*-element sequence from 88 of the 88 resistant strains (DI and DR strains), 66 of these 88 with full length elements, Fisher's Exact Test on DI & DR *vs*. DS strains; $p < 2.2 \times 10^{-16}$ for the presence of either full or partial copies). Sanger sequencing of PCR product from a subset of strains containing both full-length and partial elements confirms they contain sequence corresponding to the *P*-element. The remaining 16 strains from which we were unable to amplify the *P*-element may have copies too degenerate to be detected with PCR. In fact, our DR strains probably encompass a range of types, from those with no full-length *P*-elements but some partial, suppressive copies, to weak DI-types with full length, but few or poorly expressed copies. These types would be analogous to the range of intermediate cytotypes also seen in *D. melanogaster* [*e.g,* the M', KP, Q and weak P-types of *(2,4,16)*].

Based on these three lines of evidence, we conclude that the *P*-element is the major cause of the dysgenesis we see in *D. simulans*. In fact, the *P*-element is known to be active and cause dysgenesis in *D. simulans* when artificially introduced (37).

The *P*-element has spread worldwide in *Drosophila simulans*

We reasoned that, as in *D. melanogaster*, the M and P strains might show temporal and/or geographical differences (18,19). We therefore surveyed for the presence or absence of the *P*-element in a collection of 573 lines sampled from 54 *D. simulans* populations from across the world between 1984 and 2014 [Supplementary Table 1, Figures 3 and 4; (31,32,38,39)]. We assayed these strains for the presence the *P*-element using PCR, using repeated independent amplifications and multiple sets of primers (see Materials and Methods). Strains were considered to have only partial copies if at least one exon consistently failed to amplify, usually corresponding to internal deletions, consistent with a mechanism of production by imperfect excision (15); strains with full-length elements usually also contained other, partial copies with similar internal deletions. We find that the fraction of strains containing the *P*-element increases over time on all continents surveyed. That is, in collections from 2004 to 2009, strains free of the *P*-element predominate; after 2009, most strains harbour some form of *P*-element, and many have full-length elements (Figure 2 & 3). We found a significant association between the year the strain was collected and whether it contains any copies of the *P*-element for Africa and the Americas (Figure 3, binomial GLM; Africa: $z = 4.681$, $p = 1.17 \times 10^{-6}$, North America: $z = 5.643$, $p = 1.67 \times 10^{-8}$, Europe: z value = 8.20, $p = 2.41 \times 10^{-16}$). By 2013, almost all sampled strains contained *P*-elements in some form (Figure 3). It is possible that the appearance of a spread is due to the loss of the *P*-element in strains maintained in the

lab over time; arguing against this possibility, flies taken directly from the wild also show this pattern (*i.e.*, the Portuguese flies collected in 2006 and 2013; Figure 3).

Compared to *D. melanogaster*, the spread of the *P*-element in *D. simulans* is surprisingly fast. In *D. melanogaster*, the first *P*-type fly was collected from the wild in 1954, and most wild strains were infected by 1974. Here, we detect the first evidence of the *P*-element in *D. simulans* samples from 2004, with *P*-elements appearing in nearly all lines sampled since 2013, nine years later (apart from one sample from Ethiopia which was *P*-element free; Figure 3, Figure 4, Supplementary Table 1). A similar contrast may exist between these two species for the intracellular parasite, *Wolbachia*, which spread over the course of a few years in *D. simulans* (31,38,40), and over several decades in *D. melanogaster* (41). One reason may be that populations are more stable in *D. melanogaster*, which can over-winter in temperate regions, whereas *D. simulans* must reinvade the northern limits of its range annually (42,43). In contrast to *D. melanogaster* and *D. simulans*, their close relative, *D. sechellia* remains *P*-element free as of 2014 (Figure 2B). In both species invaded by the *P*-element, selection to mediate its worst effects appears to have been strong, as evidenced by suppression of the dysgenesis phenotype in flies collected during the latter phases of the invasions (present study, (18,33)), underscoring the evolutionary challenge posed by selfish elements.

Materials and Methods

Fly strains. We used *D. simulans* isofemale lines collected across multiple locations over 15 years (see Supplementary Table 1 for details). Flies were reared on molasses–yeast–agar *Drosophila* medium at 18°C unless otherwise stated.

Hybrid dysgenesis assays. To assay for hybrid dysgenesis, we performed reciprocal crosses between lines, using 5 virgin males and females from each line at 29°C. We dissected 30 F1 female offspring (also reared at 29°C) from each cross, and assessed them for the presence or absence of two normal ovaries (scoring females lacking one or both ovaries as dysgenic, following (24)). We tested for a difference between reciprocal crosses in the number of dysgenic female offspring using Fisher's exact tests, allowing us to roughly categorize the strains as DI (dysgenesis inducing), DS (dysgenesis susceptible), or DR (dysgenesis resistant). Some heterogeneity occurred within strains; and some strains could not be unambiguously classified (4 borderline DI/DR types, and 2 borderline DR/DS types, and 3 that tested as DI and DS). These strains were not used in the analysis; in all, 177 strains were unambiguously cytotyped.

Sterility Assays. To assay for sterility, we performed reciprocal crosses as described previously at 29°C, and compared the fertility of females collected from the dysgenic and non-dysgenic directions of the same cross. We crossed individual virgin females and control virgin males (raised at 25°C) at 25°C. We also tested for an effect of age on fertility (inspired by (44)) , by using either young (3-9 days old) or old (10-16 days) females for these crosses. We removed parents after 7 days and compared the number of offspring between the offspring of reciprocal crosses using a Mann-Whitney U test (see supplementary experimental protocols for more detail).

Analysis of RNAseq data. Paired end Illumina reads (accession number PRJEB7936 (23)) were mapped to a *D. simulans* reference genome produced from a Madagascar strain (36) and 179 *Drosophila* transposable element (TE) sequences (http://flybase.org/; (45,46)) using GSNAP (parameters: -n =1, -N = 1)(47).

Analysis of whole genome sequence data. We used paired-end Illumina reads (SRA PRJEB7936; (23)) collected from 13 barcoded individual F1 offspring of the M252 reference strain females crossed to males from Florida (2010) isofemale lines, and the paired end reads from the M252 reference strain (36). We mapped these data to the *D. simulans* M252 reference genome using BWA-SW (default parameters) (48–50) and following the mapping protocol described in Kofler et al (2012). We counted the number of insertions per individual sequenced in two ways: First, we used PoPoolationTE (50), with the requirement that each insertion be supported by at least 3 reads. Second, due to the fragmented nature of the *D. simulans* genome sequence, some insertion sites may not be recovered by PoPoolationTE, which relies on mapping reads to a reference. We therefore also estimated TE copy number per family by comparing average coverage of the TE sequence to average coverage of chromosome arm 2L.

PCR. Primers for TEs were designed based on the EMBL sequences (Supplementary Table 6, http://flybase.org/; (45,46)). The following PCR program was used: 94°C for 5' followed by 30 cycles of 94°C 30" , 55°C for 60", 72°C for 1-3' (depending on the expected length of the PCR product, 1 minute per kilobase); 72°C for 10'. PCR products were run on a 1% agarose gel to test for the presence or absence of a TE. To survey for the *P*-element, primers were designed for each exon of *P*-element (Table S3).

The presence of full *P*-elements was assessed using the forward primer for exon 0 and the reverse primer for exon 3 and using the TIR primer. *D. melanogaster* Harwich P was used as a positive control in all PCRs. Sanger sequencing of a subset of PCR products was used to confirm that the amplicons correspond to *P*-element sequence; the sequence of additional sequencing primers is available upon request.

RT-PCR. RNA was isolated from Florida, Georgia, Harvard, Maryland and Madagascar lines using Peqlab TriFast RNA isolation protocol, and extractions checked on a 1% agarose gel. We then performed RT-PCR using QIAGEN OneStep RT-PCR Kit, using the following program to detect full-length *P*-element RNA (using the forward primer for exon 0 and the reverse primer for exon 3): 45°C for 30' for reverse transcription, then 95C for 15', followed by 30 cycles of 94C for 15", 55C for 60" and 68C for 6'; 68C for 10' to amplify cDNA.

Acknowledgements

This project was supported by FWF grant P27048 to AJB. The Florida DNA sequence data & RNA sequence data is available on the NCBI sequence read archive (PRJEB7936). The Madagascar DNA sequence data is also available on the NCBI sequence read archive (SRX504933).

We report no conflicts of interest for this project.

Thanks to Viola Nolte for assistance with PCR, performing the sequencing and for providing information for many of the isofemale lines. We are grateful for helpful discussion provided by B. Charlesworth, R. Kofler, G. Lee, W. Miller, and C. Vogl. We are especially grateful to all the researchers who generously shared fly collections (Table S1), as this work would not have been possible without them.

References

1. McClintock B. Induction of instability at selected loci in Maize. Genetics. 1953;38(6):579–99.

2. Orgel LE, Crick FHC. Selfish DNA: the ultimate parasite. Nature. 1980;284:604–7.

3. Doolittle WF, Sapienza C. Selfish genes, the phenotype paradigm and genome evolution. Nature. 1980;284(5757):601–3.

4. Gregory TR. Synergy between sequence and size in large-scale genomics. Nat Rev Genet. 2005;6(9):699–708.

5. Lee YCG. The Role of piRNA-Mediated Epigenetic Silencing in the Population Dynamics of Transposable Elements in Drosophila melanogaster. PLOS Genet. 2015;11(6):1–24.

6. Pasyukova EG, Nuzhdin SV, Morozova TV, Mackay TFC. Accumulation of transposable elements in the genome of *Drosophila melanogaster* is associated with a decrease in fitness. J Hered. 2004;95(4):284–90.

7. Spradling AC, Stern D, Beaton A, Rhem EJ, Laverty T, Mozden N, et al. The Berkeley *Drosophila* genome project gene disruption project: Single P-element insertions mutating 25% of vital *Drosophila* genes. Genetics. 1999;153(1):135–77.

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

9. Hartl DL, Lozovskaya ER, Nurminsky DI, Lohe AR. What restricts the activity of mariner-like transposable elements? Trends Genet. 1997;13(5):197–201.

10. Sánchez-Gracia A, Maside X, Charlesworth B. High rate of horizontal transfer of transposable elements in Drosophila. Trends Genet. 2005;21(4):200–3.

11. Daniels SB, Clark SH, Kidwell MG, Chovnickt A. Genetic Transformation of *Drosophila melanogaster* with an Autonomous P-element: Phenotypic and Molecular Analyses of Long-Established Transformed Lines. Genetics. 1987;723:711–23.

12. Engels WR. P elements in *Drosophila*. In: Saedley H, Gierl A, editors. Transposable Elements [Internet]. Berlin: Springer-Verlag; 1996. p. 103–23. Available from: http://engels.genetics.wisc.edu/Pelements/Pt.html

13. Kidwell JF, Kidwell MG, Sved JA. Hybrid dysgenesis in *Drosophila melanogaster*: A syndrome of aberrant traits including mutation, sterility, and male recombination. Genetics. 1977;86:813–33.

14. Hiraizumi Y. Spontaneous recombination in *Drosophila melanogaster* males. Proc Natl Acad Sci U S A. 1971 Feb;68(2):268–70.

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

16. Engels WR. The P Family of Transposable Elements in *Drosophila*. Annu Rev Genet. 1983;17:315–44.

17. Daniels SB, Peterson KR, Strausbaugh LD, Kidwell MG, Chovnick A. Evidence for horizontal transmission of the P transposable element between *Drosophila* species. Genetics. 1990;124:339–55.

18. Anxolabéhère D, Kidwell MG, Periquet G. Molecular Characteristics of Diverse Populations Are Consistent with the Hypothesis of a Recent Invasion of *Drosophila melanogaster* by Mobile P Elements. Mol Biol Evol. 1988;5(3):252–69.

19. Anxolabéhère D, Hu K, Nouaud D, Periquet G. The distribution of the P-M system in *Drosophila melanogaster* strains from the People's Republic of China. Genet Sel Evol. 1990;22:175–88.

20. Daniels SB, Strausbaugh LD. The distribution of P-element sequences in *Drosophila*: The *willistoni* and *saltans* species groups. 1986;23(2):138–48.

21. Rubin GM, Spradling AC. Genetic transformation of Drosophila with transposable element vectors. Science. 1982;218:348–53.

22. Brookfield JFY. Apparent absence of transposable elements related to the P-elements of *D. melanogaster* in other species of *Drosophila*. Nature. 1984;310:330–2.

23. Kofler R, Hill T, Nolte V, Betancourt AJ, Schlötterer C. The recent invasion of natural Drosophila simulans populations by the P-element. Proc Natl Acad Sci. 2015;112(21):6659–63.

24. Engels WR, Preston CR. Hybrid Dysgenesis in *Drosophila melanogaster*: The Biology of Female and Male Sterility. Genetics. 1979;92(2257):161–74.

25. Aravin AA, Hannon GJ, Brennecke J. The piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science. 2007;318:761–4.

26. Blackman RK, Grimaila R, Koehler MMD, Gelbart WM. Mobilization of hobo Elements Residing within the Decapentaplegic Gene Complex: Suggestion of a New Hybrid Dysgenesis System in *Drosophila melanogaster*. Cell. 1987;49:497–505.

27. Bucheton A, Paro R, Sang HM, Pelisson A, Finnegan DJ. The molecular basis of I-R hybrid Dysgenesis in *Drosophila melanogaster*: Identification, cloning, and properties of the I factor. Cell. 1984;38(1):153–63.

28. Evgen'ev MB, Zelentsova H, Shostak N, Kozitsina M, Barskyi V, Lankenau D, et al. Penelope , a new family of transposable elements and its possible role in hybrid dysgenesis in *Drosophila virilis*. Proc Natl Acad Sci U S A. 1997;94(January):196–201.

29. Verspoor RL, Haddrill PR. Genetic diversity, population structure and *Wolbachia* infection status in a worldwide sample of *Drosophila melanogaster* and *D. simulans* populations. PloS One. 2011 Jan;6(10):e26318.

30. Kidwell JF, Kidwell MG, Sved JA. Hybrid dysgenesis in *Drosophila melanogaster*: A syndrome of aberrant traits including mutation, sterility, and male recombination. Genetics. 1977;86:813–33.

31. Kriesner P, Hoffmann AA, Lee SF, Turelli M, Weeks AR. Rapid Sequential Spread of Two *Wolbachia* Variants in *Drosophila simulans*. PLoS Pathog. 2013;9(9):e1003607.

32. Rogers RL, Cridland JM, Shao L, Hu TT, Andolfatto P, Thornton KR. Landscape of Standing Variation for Tandem Duplications in Drosophila yakuba and Drosophila simulans. Mol Biol Evol. 2014 Jul;31(7):1750–66.

33. Black DM, Jackson MS, Kidwell MG, Dover GA. KP elements repress P-induced hybrid dysgenesis in *Drosophila melanogaster*. EMBO. 1987;6(13):4125–35.

34. Lu J, Clark AG. Population dynamics of PIWI-interacting RNAs (piRNAs) and their targets in *Drosophila*. Genome Res. 2010;20:212–27.

35. Malone CD, Hannon GJ. Molecular evolution of piRNA and transposon control pathways in *Drosophila*. Cold Spring Harb Symp Quant Biol. 2009;74:225–34.

36. Palmieri N, Nolte V, Chen J, Schlötterer C. Genome assembly and annotation of a *Drosophila simulans* strain from Madagascar. Mol Ecol Resour. 2014;

37. Daniels S, Chovnick A, Kidwell M. Hybrid dysgenesis in Drosophila simulans lines transformed with autonomous P elements. Genetics. 121:281–91.

38. Turelli M, Hoffman A. Rapid spread of an inherited incompatibility factor in California Drosophila. Nature. 353:440–2.

39. Pool J, Langley CH. DPGP3 [Internet]. 2013. Available from: www.dpgp.org/dpgp3

40. Zhou W, Rousset F, O'Neill S. Phylogeny and PCR-based classification of *Wolbachia* strains using wsp gene sequences. Proc R Soc. 1998;265:509–15.

41. Riegler M, Sidhu M, Miller W, O'Neill S. Evidence for a global Wolbachia replacement in Drosophila melanogaster. Curr Biol. 2005;15(15):1428–33.

42. Hoffman A. Physiological climatic limits in Drosophila: patterns and implications. J Exp Biol. 2010;213(6):870–80.

43. Gravot E, Huet M, Veuille M. Effect of breeding structure on population genetic parameters in *Drosophila*. Genetics. 2004 02;166(2):779–88.

44. Khurana JS, Wang J, Xu J, Koppetsch BS, Thomson TC, Nowosielska A, et al. Adaptation to P Element transposon invasion in *Drosophila melanogaster*. Cell. 2011;147(7):1551–63.

45. Quesneville H, Bergman CM, Andrieu O, Autard D, Nouaud D, Ashburner M, et al. Combined Evidence Annotation of Transposable Elements in Genome Sequences. PLoS Comput Biol. 2005;1(2):0166–75.

46. Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, Patel S, et al. The Transposable Elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. Genome Biol. 2002;3:RESEARCH0084.

47. Wu TD, Nacu S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinforma Oxf Engl. 2010 Apr 1;26(7):873–81.

48. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinforma Oxf Engl. 2009;25(14):1754–60.

49. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinforma Oxf Engl. 2009;25(16):2078–9.

50. Kofler R, Betancourt AJ, Schlötterer C. Sequencing of pooled DNA Samples (Pool-Seq) uncovers complex dynamics of transposable element insertions in *Drosophila melanogaster*. PloS Genet. 2012;8(1):1–16.

51. Braig HR, Zhou W, Dobson SL, Neill SLO. Cloning and Characterization of a Gene Encoding the Major Surface Protein of the Bacterial Endosymbiont *Wolbachia pipientis*. J Bacteriol. 1998;180(9):2373–8.

Figure 1. Hybrid dysgenesis in *Drosophila.* Figure illustrating a normal and dysgenic cross in *Drosophila*. In the non-dysgenic cross (left), *P*-element suppressing piRNAs are deposited in the zygote from the maternal line, producing offspring with normal ovaries. In the dysgenic cross (right), the maternal line does not produce these piRNAs, none are deposited in the egg, and the offspring can suffer from hybrid dysgenesis due to unsuppressed transposition.

Figure 2. Cytotypes of lines assayed for dysgenesis. The cytotypes and *P*-element infection status of a worldwide sample of *D. simulans* lines is shown. Populations from the same year are grouped together by cytotype; see Figure 3 or Supplementary Table 1 for locations. Two strains showed amplification of all four exons individually, but not of full-length element.

Figure 3. Partial and full *P*-element copies in *D. simulans* collected A) 1984-2008, and B) 2009-2014. Map shows the approximate location where strains were collected, and pie charts show the proportion of strains containing *P*-element (either full-length copies or individual exons could be amplified by PCR) and no copies (no *P*-element of any kind could be sampled). The area of the pie charts is proportional to the number of strains sampled (see Legend), except for single points, which indicate one strain.

Figure 4. Time series of *P*-element in different forms in *D. simulans*. The plot shows strains with full (and partial) and no *P*-elements over time in Africa (top panel), North and South America (middle panel), and Europe (bottom panel) across

time. Each point indicates a strain, with overlapping points jittered slightly. Data from the 1980's in the Americas are extracted from Brookfield (1984) (22).

Figure 1.

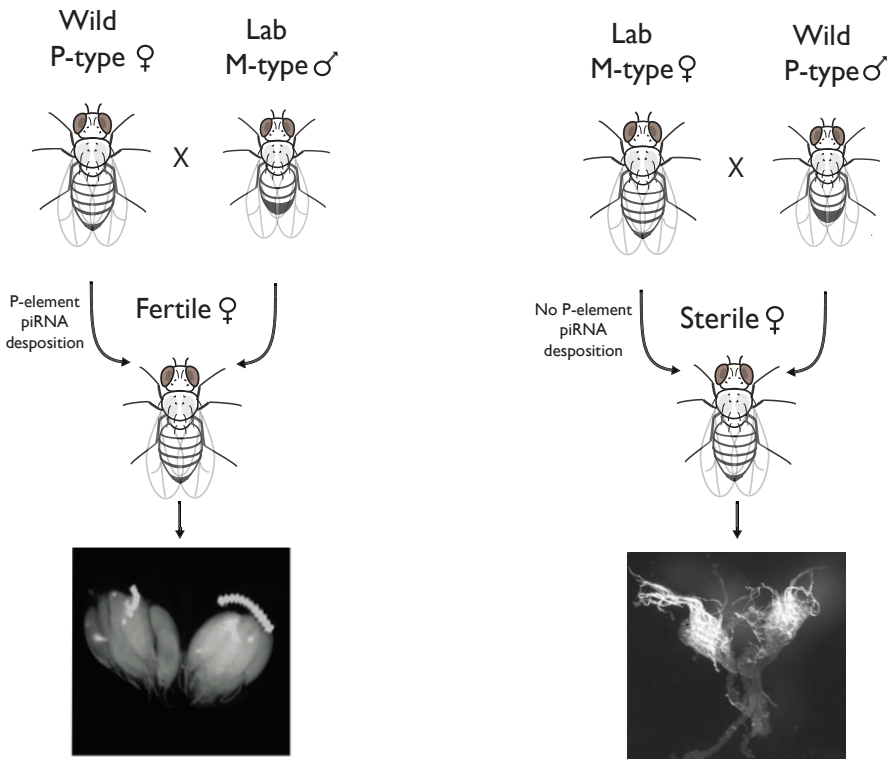

Wild
P-type ♀
Lab
M-type ♂
X
P-element
piRNA
desposition
Fertile ♀
Lab
M-type ♀
Wild
P-type ♂
X
No P-element
piRNA
desposition
Sterile ♀

Figure 2.

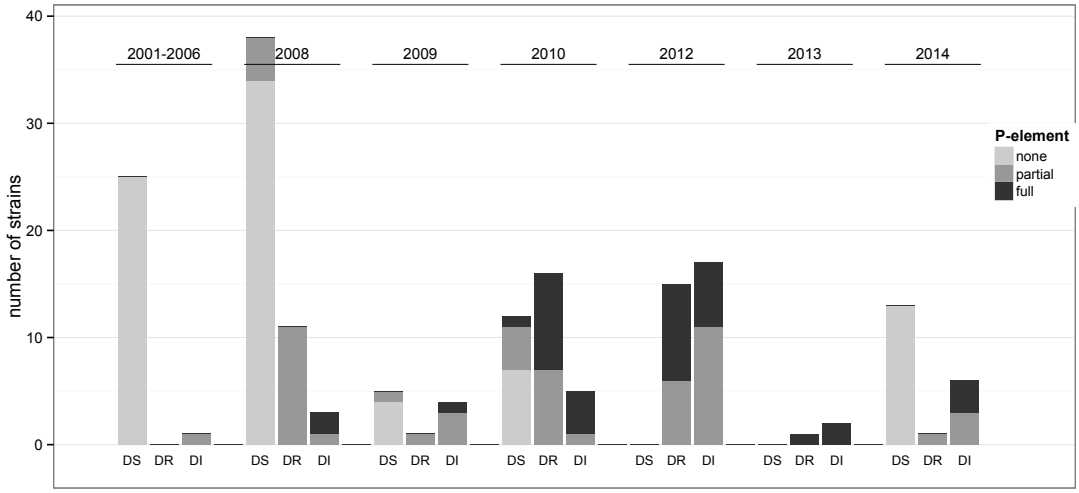
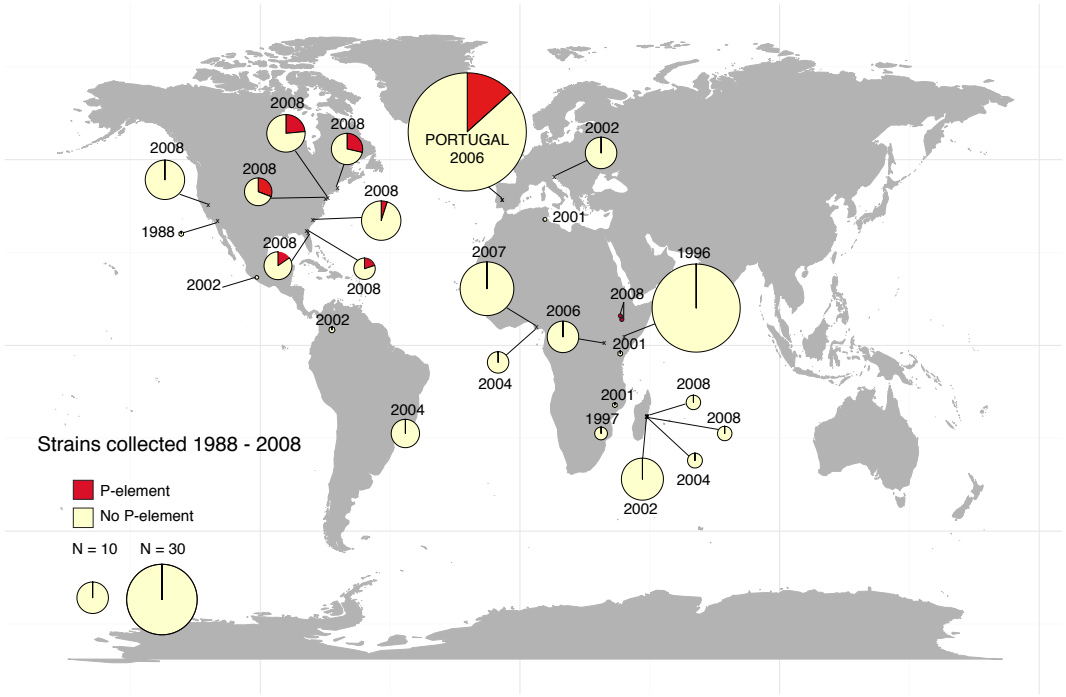

40
30
20
10
0
number of strains
2001-2006
2008
2009
2010
2012
2013
2014
DS DR DI
P-element
none
partial
full

Figure 3.

A


PORTUGAL
2006
2008
2008
2008
2008
2008
2008
2008
1988
2002
2002
2002
2001
2007
2006
2004
2004
1996
2008
2001
2001
1997
2008
2008
2004
2002
Strains collected 1988 - 2008
P-element
No P-element
N = 10
N = 30

B

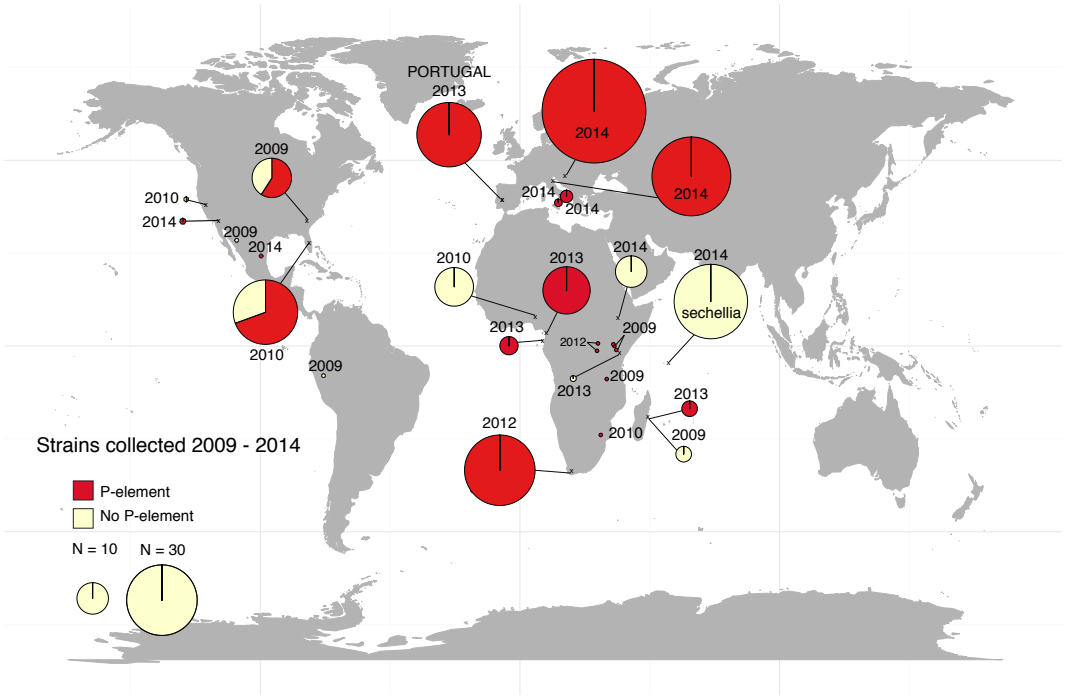
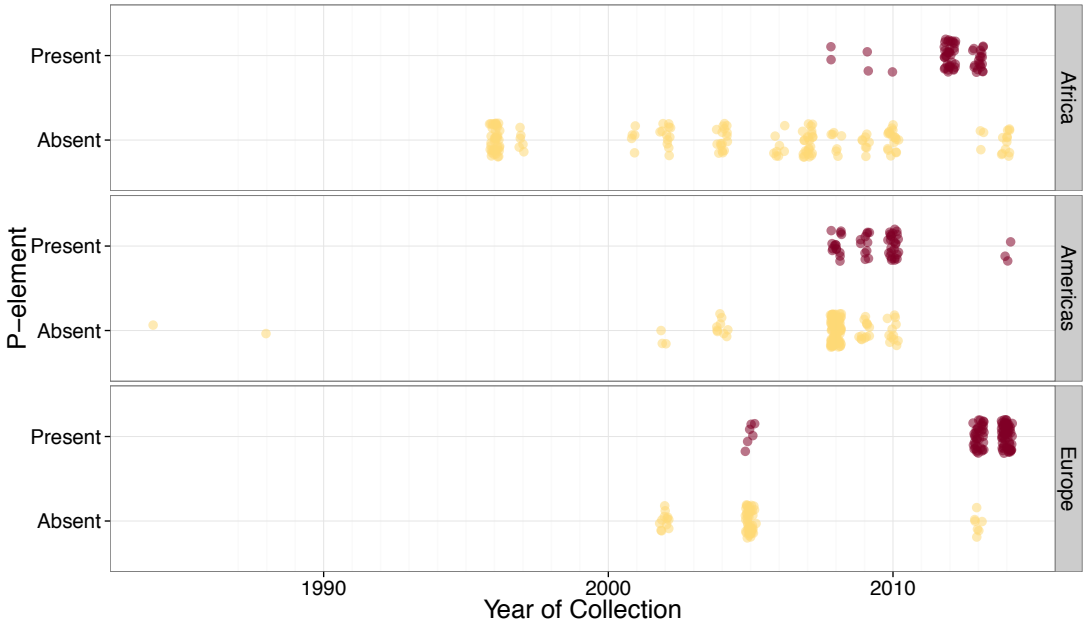

PORTUGAL
2013
2014
2014
2014
2014
2009
2010
2014
2009
2014
2010
2010
2013
2014
2014
sechellia
2013
2012
2009
2009
2009
2013
2013
2012
2010
2009
Strains collected 2009 - 2014
P-element
No P-element
N = 10
N = 30

Figure 4.


Present
Absent
Africa
Present
Absent
Americas
Present
Absent
Europe
P−element
1990
2000
2010
Year of Collection