Supplementary Material and Methods

Experimental populations

We established 202 isofemale lines from a *D. simulans* population collected 2010 (November) in Florida (Tallahassee) and maintained them in the lab for 6-7 generations at 18°C. To set up base populations (Spring 2011) we mixed 5 mated females from each isofemale line (N=1010) and let them lay eggs twice, for 24h each. These two sets of progeny were used to establish the hot and cold evolving populations (see below). We repeated this procedure (mixing 5 females from each isofemale line, let them lay eggs twice) to generate 10 replicates, of hot as well as cold evolving populations (H1, H2,..., H10; C1, C2,..., C10). Note that this procedure generates pairs of populations having exactly the same base populations (H1 and C1, H2 and C2,..., H10 and C10).

We kept each replicate at a census population size of about 1000 in 5 bottles (6 oz) containing 50 mL standard Drosophila medium. To establish a new generation we collected all adults approximately 3 days after eclosion, mixed them, distributed 1000 randomly picked flies equally to 5 new bottles and allowed them to lay eggs for 48h. After egg laying the adults were either discarded or used to monitor the status of the invasion (e.g.: Pool-seq, RNA-seq, small RNA sequencing, hybrid dysgenesis; see below). To mimic the natural day-night cycle of Drosophila the temperature cycled between two fixed temperatures and illuminations every 12h. At hot conditions the temperature cycled between 18°C and 28°C and at cold conditions between 10°C and 20°C. Superimposed with the cycling of the temperature, the illumination alternated between light and dark, where the light condition overlapped with the high temperature phase and the dark condition with the low temperature phase.

Genome sequencing

To monitor the abundance of the P-element in the evolving populations we sequenced the populations each 10 generations as pools [Pool-seq; (Schlötterer et al., 2014)]. We extracted genomic DNA from flies aged 6-8 days using a high salt extraction protocol (Miller et al., 1988) and sheared the DNA with a Covaris S2 device (Covaris, Inc. Woburn, MA, USA). Paired-end libraries were prepared using the TruSeq DNA PCR-Free protocol (Illumina, San Diego, CA). We sequenced the samples using the Illumina paired-end technology and trimmed all reads to a length of 100bp (to avoid biases). For an overview of the sequenced samples see supplementary table 1.

Estimating TE abundance using Pool-seq data

We estimated the TE abundance in the Pool-seq libraries with PoPoolationTE2 (v1.10.04) (Kofler et al., 2016) which requires a reference genome, consensus sequences of TE families and a TE hierarchy. As reference genome we used the sequence of strain strain M252 (Palmieri et al., 2015) (GenBank ID JMCE00000000; we included the sequences of *Lactobacillus brevis*, *Acetobacter pasteurianus* and two Wolbachia strains; GenBank accession

Supplementary Table 1: Overview of the Pool-seq data used for estimating the status of the P-element invasion; We show the IDs of the used lanes, the read length (rl.), the insert size (is.) and the number of sequenced paired end reads (in millions) for each replicate. Reads of samples 185a,b,c and 238a were trimmed to a length of 100bp. ts-r1: time series for hot replicate 1

					replicate		
sample	lane IDs	Sequencer	rl.	is.	1	3	5
base	126a,b,c,d,e	HiSeq 2000	100	300	73.5	105.8	122.1
hot10	127a,b	HiSeq 2000	100	380	34.1	27.0	44.7
hot20	128a,b,c,d	HiSeq 2000	100	330	88.0	85.0	79.0
hot30	129b,c,d	HiSeq 2000	100	330	40.8	36.5	22.2
hot40	130b,c,d	HiSeq 2000	100	330	59.0	75.9	50.4
hot50	131b,c	HiSeq 2000	100	330	29.3	30.0	36.5
hot60	140b,c	HiSeq 2000	100	380	48.6	27.9	31.6
cold10	132b,c	HiSeq 2000	100	330	34.7	40.6	34.2
cold20	133b,c	HiSeq 2000	100	340	42.7	25.0	24.4
cold30	134b,c	HiSeq 2000	100	340	29.2	32.7	33.1
cold40	185a,b,c	HiSeq 2500	120 (100)	530	19.4	46.9	21.7
ts-r1	238a	HiSeq 2500	125 (100)	530	224.1	-	-

numbers CP000416.1, AP011170.1, AE017196.1, CP001391.1) and obtained a set of consensus TE sequences from FlyBase (http://flybase.org; transposon_sequence_set.embl; v9.42; (Quesneville et al., 2005)). We obtained the TE hierarchy from the database of consensus TE sequences and annotated TEs in the reference genome as described before (Kofler et al., 2015b).

We extracted the sequences of annotated TE insertions from the reference genome into a distinct file and subsequently masked these TE sequences in the reference genome using the character N. Next we concatenated the fasta records of (i) the consensus sequences of TE insertions, (ii) the TE sequences extracted from the reference genome and (iii) the repeat masked reference genome into a single file (the TE-merged-reference). We mapped all reads to the TE-merged-reference using bwa bwasw (v0.7.5a) (Li and Durbin, 2010) and restored paired-end information using PoPoolationTE [*samro.pl*; (Kofler et al., 2012)]. Basic mapping statistics, including the fraction of reads mapping to each TE family (in rpm; reads per million), were computed with PoPoolationTE2 [*stat-reads* (Kofler et al., 2016)] and samtools (v.1.2) (Li et al., 2009). Using PoPoolationTE2 [v1.10.04 (Kofler et al., 2016)] we generated a ppileup (physical pileup) file (*ppileup* –map-qual 15), subsampled the ppileup to a uniform coverage of 15 (*subsamplePpileup* –target-coverage 15), identified signatures of TE insertions (*identifySignatures* –mode separate, –min-count 1, –signature-window fix100), identified the strand of TE signatures (*updatestrand* –map-qual 15, –max-disagreement 0.4), estimated the population frequencies of TE signatures (*frequency*), filtered TE signatures

(*filterSignatures* –min-coverage 10, –max-otherte-count 2) and finally paired TE signatures to obtain a list of all identified TE insertions (*pairupSignatures* –min-distance -200 –max-distance 300).

The average number of P-element insertions per diploid genome was estimated using the equation: $counts = 2 * rpm * G/L$ where rpm is the fraction of reads mapping to the P-element (in reads per million), $2 * G$ the diploid genome size of *D. simulans* in Mb [$2 * 162$; Bosco et al. (2007)] and L the length of the P-element [2907; O'Hare and Rubin (1983)].

The enrichment of P-element insertions in origin recognition complex (ORC) binding sites was computed as described before (Kofler et al., 2015a). Briefly we obtained ORC positions for *D. melanogaster* from http://flypush.imgen.bcm.tmc.edu/pscreen/downloads.html (Spradling et al., 2011), merged overlapping ORCs (bedtools merge v2.26.0; Quinlan and Hall (2010)), extracted the ORC sequences from the reference genome of *D. melanogaster* (bedtools getfasta v2.26.0; *D. melanogaster* v5.53), mapped these sequences to the *D. simulans* reference genome with bwa bwasw (v0.7.15; Li and Durbin (2010)), filtered for a minimum length of 100bp, merged overlapping ORCs (bedtools merge, v2.26.0) and identified P-element insertions overlapping with an ORC (bedtools intersect; v2.26.0). To identify the fraction of P-elements inserted in another TE we obtained all paired ends reads where the first read maps to the P-element and the second read does not map to the P-element. Finally we estimated the fraction of fragments with the second read mapping to another TE (not the P-element). We performed this analysis for all reads from hot and cold conditions separately.

Estimating the abundance of P-elements with internal deletions

The break points of internally deleted P-elements may be identified by mapping reads with an algorithm that allows for large indels [see for example Kofler et al. (2015a)]. We first filtered for reads mapping to the P-element with samtools [v.1.2;(Li et al., 2009)], converted the mapped reads to a fastq file (*sam2fastq.py*) and realigned the reads to the P-element using gsnap [v2014-10-22; –novelsplicing=1; gsnap is designed for RNA-seq data and thus allows for large indels resulting from spliced introns (Wu and Nacu, 2010)]. Finally we estimated the frequency of internal deletions as the ratio of the number of reads supporting a given deletion to the average coverage of the P-element, with deleted regions contributing to the coverage (*trunc-freq-from-bam.py*). Note that this approach does not allow to estimate the genomic position of internally deleted P-elements.

To generate "fitness landscapes" of internally deleted P-elements we computed for each site of the P-element the average frequency of all internal deletions spanning the given site. For hot evolved populations we used the frequency at generation 60 and for cold evolved flies the frequency at generation 40. For internal deletions occurring in more than one replicate the largest frequency was used (the fitness landscape is computed across replicates). Fitness landscapes were computed with a custom script (*generate-siteenrichment-avfreq.py*)

To obtain unbiased estimates of the abundance of internally deleted P-elements in hot and cold evolved populations we subsampled reads mapped to the P-element with gsnap

(see above) to a final coverage of 52 (with internal deletions contributing to the coverage). We repeated this procedure 100 times for each sample and estimated the mean number of identified internal deletions and the standard deviation.

Sashimi plots providing an overview of the position and abundance of internally deleted P-elements were generated as described previously (Kofler et al., 2015a).

Polymorphism of the P-element

We generated a mpileup-file [samtools v1.2; -B -Q 0; (Li et al., 2009)] using reads mapping to the P-element with gsnap (see section 2). The mpileup file was converted to a sync file with PoPoolation2 [r204; –fastq-type sanger –min-qual 20 (Kofler et al., 2011)] and SNPs were identified with a custom tool (*find-snps.py*) where we required a minor allele count of 20 across all samples (of a replicate) for hot evolved populations and 5 for cold evolved populations.

RNA-seq

We estimated the expression of the P-element in replicates 8, 9 and 10 of hot (generation 22) and cold evolved (generation 11) populations using common gardens. Flies from both the hot and cold evolved populations were allowed to lay eggs twice for 48 hours. These eggs were kept in a hot common garden (constant 23° C) and in a cold common garden (constant 15° C) for 2 generations. At the last generation we performed a density control by randomly selecting 300 eggs. This yields in total 12 samples (3 replicates, 2 temperature conditions and 2 common gardens). Finally, 50 virgin females (hot: aged 3-4 days; cold: aged 6-8 days) were randomly picked from each sample and snapfrozen in liquid nitrogen. Total RNA was extracted using TriFast (Peqlab, Erlangen, Germany), subjected to DNase I treatment and enriched for polyA-mRNA using oligo(dT) magnetic beads. Using a modified NEB Next Ultra protocol we generated strand-specific paired-end libraries with an insert size of about 300bp. We barcoded the samples and sequenced three lanes of 2x100bp paired-end reads on a Illumina HiSeq 2000 (Illumina, San Diego, CA). The reads were mapped to the TE-merged-reference (see section 2) - which contains the consensus sequence of the P-element - using bwa bwasw (v0.7.5a) (Li and Durbin, 2010). The fraction of reads mapping to the P-element was computed with a custom script (*readstat2rpm.py*). For an overview of the RNA-seq data used in this work see table 2. Significant differences in P-element expression of flies reared in a hot and cold common garden were assessed with an ANOVA. We compared a model that includes both the temperature of the cage and the common garden to a model that only includes the temperature of the cage as explanatory variable.

To identify alternative splicing we realigned all reads mapping to the P-element with gsnap (section 2) and counted for each splice junction the number of reads supporting the presence or absence of a splicing event (*delta23freq-from-bam.py*). For each intron we summed values of the 5' and 3' splice junction.

Supplementary Table 2: Overview of the RNA-seq libraries used in this work. For each library the number of reads (in million: M.), the fraction of mapped reads, the number of reads mapping to the P-element (P [counts]) and the fraction of reads mapping to the P-element (P [rpm], reads per million) is shown. rep.: replicate, cg.: common garden

rep.	cg.	reads [M.]	mapped [%]	P [counts]	P [rpm]
H8	hot	52.4	98.3	2260	43.9
	cold	50.8	99.1	1056	21.0
H9	hot	64.8	99.3	4069	63.2
	cold	42.2	99.3	470	11.2
H10	hot	50.2	98.9	1874	37.8
	cold	41.8	99.2	764	18.4
C8	hot	50.0	99.0	229	4.6
	cold	50.3	99.2	132	2.6
C9	hot	69.8	99.3	372	5.4
	cold	48.0	99.3	94	2.0
C10	hot	42.7	99.1	91	2.1
	cold	48.4	99.2	34	0.7

Small RNAs

For hot evolved populations, small RNA was sampled at generations 22, 44 and 108; and for cold evolved populations at generations 22 and 54. For an overview of the small RNA data see supplementary table 5. Total RNA was extracted from females (aged 3-4 days) reared in a hot common garden hot (constant 23°C) as described above (section 2). Sequencing of the small RNA libraries was done by Fasteris (https://www.fasteris.com/dna/). We mapped the short reads with novoalign (V3.03.02; http://www.novocraft.com/ -F STDFQ -o SAM -o FullNW -r RANDOM) to a library consisting of the *D. simulans* tRNAs, miRNAs, mRNAs, snRNAs, snoRNAs, rRNAs (1.4; Flybase http://flybase.org/) and the consensus sequences of the TEs (transposon_sequence_set.embl; v9.42; section 2). An overview of the mapping results, the distribution of piRNAs within the P-element (including normalization to a million miRNA) and the ping-pong signal were computed with custom Python scripts (*stat-piRNA-overview.py*, *graph-piRNA-distribution-onTE.py*, *ping-pong-signature.py*; allowing for 2 mismatches). Small RNAs from the *D. melanogaster* strain Harwich were obtained from ENA [http://www.ebi.ac.uk/ena; accession number SRR014273; published by Brennecke et al. (2008)]. Adapters were removed with cutadapt [v1.9.1 (Martin, 2011); -a CTGTAGGCACCATCAAT] and the sequences were aligned with novoalign as described above. We identified the positions of piRNA clusters for hot and cold conditions separately. We pooled the small RNA data from generation 22 for each temperature regime and mapped them to the *D. simulans* genome. To exclude siRNAs and miRNas we only retained reads with a length between 24 and 29bp. Furthermore solely unambiguously mapping reads were

used (mapping quality ≥ 20). Next we counted the number of piRNAs for genomic bins of size 100bp ($(bin - pirnas.py - -binsize100$) and than identified the positions of piRNA clusters using a local score approach as described by Fariello et al. (2017). Briefly we computed a score for every bin (number of piRNAs - threshold) and searched for sets of consecutive bins that attain a local high score. Note that this algorithm allows for bins having negative scores as long as a novel high score is achieved with the neighboring bins ($findclusterinsertions.py - -threshold50$). Finally we only retained piRNA clusters with a local score larger than 2000. To test whether the P-element insertions on 3R are close to telomere associated regions (TAS) we obtained the TLL (TAS-L like) sequences of *D. simulans* provided (Asif-Laidin et al., 2017) (Supplementary fig. 7A,B) and aligned them to the *D. simulans* genome with novoalign (-F FA -o SAM -r All).

Single molecule RNA-FISH

We determined the presence of P-element transcripts in hot and cold evolved populations (hot: generation 108; cold: generation: 54) using single molecule RNA-FISH. Flies were reared for two generations in a hot common garden (constant 23° C) as described above (section 2). We used CAL Fluor Red 590-labeled Stellaris oligo probes (LGC Biosearch Technologies, Petaluma, CA) to detect P-element transcripts in *D. simulans* ovaries. We designed one probe set for sense mRNA and two probe sets for antisense mRNA of the P-element. For each probe set, 48 oligos were designed such that each oligo has at least two mismatches to any other TE. Single molecule RNA-FISH was performed as described in Mohn et al. (2014). Confocal stacks of egg chambers were acquired on a Zeiss LSM780 microscope.

Gonadal dysgenesis assays

We measured the abundance of dysgenic ovaries in hot and cold evolved populations (hot: generation 114; cold: generation 57). Flies were randomly assigned to three samples and each sample was allowed to lay eggs for 48h at one of three temperature regimes (29°C constant; hot fluctuating 18-28° C; cold fluctuating 10-20°C). The progeny was developing at the respective temperature regime until eclosion and eclosed flies were kept at constant 23°C. To estimate the fraction of dysgenic ovaries we kept flies on apple juice agar with live yeast for 2 days. Flies were dissected in PBS and presence of dysgenic ovaries was scored using the following classification: clearly visible ovarioles or eggs (clear), ovarioles barely visible (weak), no ovarioles (or eggs) could be detected (absent). The percentage of dysgenic ovaries (HD) was computed as $100 * (absent + (weak/2))/(clear + weak + absent)$

References

Asif-Laidin, A., Delmarre, V., Laurentie, J., Miller, W. J., Ronsseray, S., and Teysset, L. (2017). Short and long-term evolutionary dynamics of subtelomeric piRNA clusters in Drosophila. *DNA Research*, 0:1–14.

Beall, E. L., Mahoney, M. B., and Rio, D. C. (2002). Identification and analysis of a hyperactive mutant form of drosophila p-element transposase. *Genetics*, 162(1):217–227.

Bellen, H. J., Levis, R. W., He, Y., Carlson, J. W., Evans-Holm, M., Bae, E., Kim, J., Metaxakis, A., Savakis, C., Schulze, K. L., Hoskins, R. A., and Spradling, A. C. (2011). The Drosophila gene disruption project: progress using transposons with distinctive site specificities. *Genetics*, 188(3):731–43.

Black, D. M., Jackson, M. S., Kidwell, M. G., and Dover, G. A. (1987). KP elements repress P-induced hybrid dysgenesis in *Drosophila melanogaster*. *The EMBO journal*, 6(13):4125–35.

Bosco, G., Campbell, P., Leiva-Neto, J. T., and Markow, T. A. (2007). Analysis of Drosophila Species Genome Size and Satellite DNA Content Reveals Significant Differences Among Strains as Well as Between Species. *Genetics*, 177(3):1277–1290.

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

Fariello, M. I., Boitard, S., Mercier, S., Robelin, D., Faraut, T., Arnould, C., Recoquillay, J., Bouchez, O., Salin, G., Dehais, P., Gourichon, D., Leroux, S., Pitel, F., Leterrier, C., and SanCristobal, M. (2017). Accounting for linkage disequilibrium in genome scans for selection without individual genotypes: The local score approach. *Molecular Ecology*, 26(14):3700–3714.

Kofler, R., Betancourt, A. J., and Schlötterer, C. (2012). Sequencing of Pooled DNA Samples (Pool-Seq) Uncovers Complex Dynamics of Transposable Element Insertions in *Drosophila melanogaster*. *PLoS genetics*, 8(1):e1002487.

Kofler, R., Gómez-Sánchez, D., and Schlötterer, C. (2016). PoPoolationTE2 : Comparative Population Genomics of Transposable Elements Using Pool-Seq. *Molecular Biology and Evolution*, 33(10):2759–2764.

Kofler, R., Hill, T., Nolte, V., Betancourt, A., and Schlötterer, C. (2015a). The recent invasion of natural *Drosophila simulans* populations by the P-element. *PNAS*, 112(21):6659–6663.

Kofler, R., Nolte, V., and Schlötterer, C. (2015b). Tempo and Mode of transposable element activity in *Drosophila*. *PLoS Genetics*, 11(7):e1005406.

Kofler, R., Pandey, R. V., and Schlötterer, C. (2011). Popoolation2: identifying differentiation between populations using sequencing of pooled dna samples (pool-seq). *Bioinformatics*, 27(24):3435–6.

Lee, C. C., Beall, E. L., and Rio, D. C. (1998). DNA binding by the KP repressor protein inhibits P-element transposase activity in vitro. *EMBO Journal*, 17(14):4166–4174.

Li, H. and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, 26(5):589–595.

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

Majumdar, S. and Rio, D. C. (2015). P transposable elements in *Drosophila melanogaster*. *Microbiol Spectrum*, 3(2):484–518.

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

Miller, S. A., Dykes, D. D., and Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic acids research*, 16(3):1215.

Mohn, F., Sienski, G., Handler, D., and Brennecke, J. (2014). The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in Drosophila. *Cell*, 157(6):1364–1379.

O'Hare, K. and Rubin, G. M. (1983). Structures of P transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. *Cell*, 34(1):25–35.

Palmieri, N., Nolte, V., Chen, J., and Schlötterer, C. (2015). Genome assembly and annotation of a *Drosophila simulans* strain from Madagascar. *Molecular ecology resources*, 15(2):372–381.

Quesneville, H., Bergman, C. M., Andrieu, O., Autard, D., Nouaud, D., Ashburner, M., and Anxolabehere, D. (2005). Combined evidence annotation of transposable elements in genome sequences. *PLoS computational biology*, 1(2):166–175.

Quinlan, A. R. and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)*, 26(6):841–842.

Robillard, É., Le Rouzic, A., Zhang, Z., Capy, P., and Hua-Van, A. (2016). Experimental evolution reveals hyperparasitic interactions among transposable elements. *Proceedings of the National Academy of Sciences*, 113(51):14763–14768.

Sandelin, A., Alkema, W., Engström, P., Wasserman, W. W., and Lenhard, B. (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Research*, 32(90001):91D–94.

Schlötterer, C., Tobler, R., Kofler, R., and Nolte, V. (2014). Sequencing pools of individuals — mining genome-wide polymorphism data without big funding. *Nature Reviews Genetics*, 15(11):749–763.

Spradling, A. C., Bellen, H. J., and Hoskins, R. A. (2011). *Drosophila* P elements preferentially transpose to replication origins. *Proceedings of the National Academy of Sciences of the United States of America*, 108(38):15948–15953.

Wu, T. D. and Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics (Oxford, England)*, 26(7):873–81.