bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

Genome Elimination Mediated by Gene Expression from a Selfish Chromosome

Authors: Elena Dalla Benetta[1,2] Igor Antoshechkin[3], Ting Yang[2], Hoa Quang My Nguyen[2], Patrick M. Ferree[1*] and Omar S. Akbari[2*]

Affiliations:

[1]W. M. Keck Science Department of Claremont McKenna, Pitzer, and Scripps Colleges, Claremont, California 91711.

[2]Division of Biological Sciences, Section of Cell and Developmental Biology, University of California, San Diego, La Jolla, California 92093.

[3]Division of Biology and Biological Engineering (BBE), California Institute of Technology, Pasadena, California 91125.

*Correspondence to: oakbari@ucsd.edu, pferree@kecksci.claremont.edu

Keywords:

Nasonia, PSR, B chromosome, Gene Drive, Selfish DNA

Abstract:

Numerous plants and animals harbor selfish B chromosomes that "drive," or transmit themselves at higher-than-Mendelian frequencies, despite long-term fitness costs to the organism. Currently it is unknown how B chromosome drive is mediated, and whether B-gene expression plays a role. We used modern sequencing technologies to analyze the fine-scale sequence composition and expression of Paternal Sex Ratio (PSR), a B chromosome in the jewel wasp *Nasonia vitripennis*. PSR causes female-to-male conversion by destroying the sperm's hereditary material in young embryos in order to drive. Using RNA interference, we show that testis-specific expression of a PSR-linked gene, named *haploidizer*, facilitates this genome elimination-and-sex conversion effect. *Haploidizer* shares homology with a gene in *Candidatus cardinium*, a bacterial symbiont that also induces genome elimination in its insect host.

One Sentence Summary:

haploidizer mediates B chromosome drive

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

Main Text:

In eukaryotes, the genome contains a core set of chromosomes that are indispensable for viability. However, thousands of plant and animal species also harbor non-essential chromosomes termed B chromosomes, consisting of both coding and non-coding sequences thought to be mostly copied from the essential, or A, chromosomes. Despite their non-essential nature, B chromosomes presumably carry no needed genes and therefore are, in principle, prone to loss over successive generations (*1–5*). To counter the tendency to be lost, many B chromosomes possess the extraordinary ability to "drive," or transmit themselves at frequencies above those predicted by Mendelian rules (reviewed in (*6, 7*)). Although drive is necessary for B chromosome transmission, it can impose deleterious effects on the inheritance of other genomic regions, thus having potentially strong influences on organismal fitness (*5, 6, 8*). Currently, it is not understood how B chromosomes can behave so differently in their transmission from other, non-driving chromosomes.

One of the most striking examples of drive is caused by a B chromosome known as PSR (for Paternal Sex Ratio), which has been detected at moderate frequencies (6-11%) within certain populations of the jewel wasp *Nasonia vitripennis (9)*. PSR is transmitted strictly paternally (*i.e.*, via sperm), and its presence leads to the complete elimination of the sperm's essential chromosomes, but not PSR itself, during the mitotic division immediately following fertilization (*10*). Genome elimination results directly from failure of the sperm's chromatin to resolve into individualized chromosomes upon entry into mitosis and is preceded by distinct chemical markers that indicate a disrupted chromatin state (*11–14*). This effect is critical for the transmission of PSR. Wasps, like all hymenopteran insects (also including ants and bees), reproduce through haplo-diploidy, in which unfertilized eggs develop into haploid males while fertilized eggs become diploid females. Therefore, elimination of the sperm's hereditary material by PSR converts all diploid, female-destined eggs into haploid males, the PSR-transmitting sex (*9*). Broadly, this effect results in severely male-biased sex ratios at the deme level, which can negatively impact wasp metapopulations (*15*).

A fundamental biological question is how PSR and, more broadly, other B chromosomes are capable of interacting uniquely with the cellular environment in order to drive. Previous studies identified three distinct ~180 DNA-base pair repeats, termed PSR2, PSR18, and PSR22, which

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

are abundant in copy number on PSR, but not present on any of the five essential wasp chromosomes (*16*, *17*). Each of these repeats contains a highly conserved, 8-bp palindromic motif that is reminiscent of those found in some transcription factor-binding sequences (*16*, *17*). These characteristics have led to speculation that PSR-specific sequences may act as a sink for some limited chromatin factor(s), drawing them away from the rest of the genome and causing defective chromatin structure (*16*). An underlying aspect of this scenario is that the driving effect of PSR is the result of intrinsic properties that are unique to its sequence composition. For this reason, we refer to such an effect as *passive*. Alternatively, PSR's drive may involve the *active* expression of B-linked sequences. Recent studies have begun to identify individual sequences that are expressed from a number of different B chromosomes, including PSR (*18*, *19*). While no studies have, to date, demonstrated functionality for any B-linked sequence, it is plausible that an RNA or protein expressed by PSR functions as an effector of drive by disrupting normal transmission of the paternal chromatin.

Previous transcriptional profiling uncovered a handful of long, polyadenylated RNAs expressed uniquely by PSR in the wasp testis (*18*). To further this work, we used modern genome sequencing technologies to generate large scaffolds of the PSR chromosome, onto which we mapped these PSR-expressed transcripts. These efforts produced a comprehensive portrayal of PSR's sequence composition, models for its individual genes, and its transcribed loci. Using RNA interference (RNAi), we show that genome elimination requires the expression of one of these PSR-linked loci, demonstrating an active (*i.e.*, gene expression-based) involvement of this B chromosome in its own drive.

To deduce PSR's sequence composition, we sequenced the genome of both wildtype and PSR-carrying wasps separately using a combination of PacBio, Oxford Nanopore and Illumina technologies, generating a 297Mb assembly comprising 444 contigs with an N50 length of 6.6 Mb representing a 13.9 fold decrease in contig number and 9.3 fold increase in N50 length compared to the current *N. vitripennis* genome assembly (*20*). Using the hymenoptera-specific set of universal single-copy orthologs, the genome completeness is estimated to be 96.5% (Table S1) (*21*). PSR-specific contigs were identified using the chromosome quotient method, which calculates the ratio of the number of wild type and PSR+ reads mapping to a contig (Table S2, Fig. S1) (*22*)) (see Methods). This approach yielded 120 contigs (9.2 MB in total, N50 of 124

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

kb) that were specific to PSR (Table S3). In addition, 73 contigs totaling 272 Mb (91.58% of total genome) were assigned to the five essential *N. vitripennis* chromosomes using the previously developed genetic markers (*23*) (Fig. S2, Table S4, Data S1.tar).

To better understand the sequence composition and organization of PSR (Table S5), we performed a detailed computational assessment of PSR's repetitive content with RepeatModeler and RepeatMasker (*24*, *25*). These analyses revealed that 89.80% of PSR is composed of repetitive DNAs (Fig.1A, Table S6). The most abundant repeats (70.32%) are complex satellites that belong to four main families. Three of these satellites, PSR2 (49.54%), PSR18 (42.64%) and PSR22 (17.22%), are specific to PSR. PSR2 and PSR18 are typically found together on gene-coding contigs and they mostly overlap, whereas PSR22 is found on different sets of contigs without coding genes. A fourth repeat, NV79 (2.71%), is located on four PSR scaffolds, not overlapping with other repeats, and also on all five essential chromosomes (Fig. 1B). Additionally, PSR contains DNA and RNA (retro-) transposable elements (TEs) (13.97%), simple repeats (0.36%) and other low complexity regions (0.03%), and uncharacterized sequences (3.81%) (Fig.1A; Table S5, S6). Two PSR contigs contain telomeric sequences that cover 3.5-3.9 kb at one of the ends of each contig. Interestingly, the telomeric repeat found on PSR and at the ends of all five A-chromosomes of *N. vitripennis* is 'TTATTGGG,' which is different from the canonical sequence 'TTAGGG' (*26*) that is found in other insect species (Fig S3).

Using Nanopore RNA sequencing from PSR-carrying testes and whole animals, we identified 68 transcripts (Table S7) encoded by 44 loci located on 14 of the 120 scaffolds and covering 8.57% of the total length of PSR (Table S8). The majority of the intronic regions are represented by TEs and low complexity regions; thus, the coding sequences corresponding to exonic regions represent only 0.51% of the entire PSR chromosome. Fifty of the PSR-encoded transcripts have identifiable open reading frames longer than 100 aa and 34 of those possess Pfam domains suggesting putative molecular functions that may be performed by these genes (Table S7). Fifty-eight transcripts produced significant blast hits when searched against the non-redundant protein database. Thirty-one sequences had matches to proteins encoded by the wild-type *N. vitripennis* genome, although *N. vitripennis* genes were the best hits for only 13 of them. The corresponding genes are located in different regions across all 5 essential chromosomes (Fig. 1C) and span a

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

range of functional groups, including transposon activity, transcription regulation, DNA binding, and protein binding. Importantly, 20 transcript sequences were best matched to genes found in the genome of a closely related parasitoid wasp, *Trichomalopsis sarcophagae*, which represents the most common lineage identified by the blast searches. The remaining 25 best matches are distributed mostly among the genomes of other insects, although other non-insect lineages, including bacteria, are represented as well (Table S8), and they mainly correspond to protein-coding regions of transposons. Thus, unlike some B chromosomes, whose sequences derived largely from one or a few large regions of essential chromosomes within their resident genomes (*27–29*), PSR consists primarily of three complex repeats (70.32%) and other sequences that are absent from the *N. vitripennis* genome and, in some cases, derived from other organisms.

Quantification of the RNAseq data revealed that 25 of the 44 PSR-encoded loci are expressed at moderate to high levels (TPM >= 10) in *Nasonia* testes (Table S8). The three highest-expressed loci were selected for further analysis (Fig. 1D, Table S8). Two of these sequences were previously identified in an earlier transcriptome profiling study as PSR4317 and PSR1539 (*18*), whereas the third highest-expressed gene, PSR-tra, has a predicted ORF with limited homology to part of the wasp's sex-determining gene, *transformer*. Based on their transcript sequences and chromosomal gene models (Fig. S4), PSR4317 and PSR1539 may contain translational open reading frames (ORFs) of 279 and 242 amino acids, respectively (*18*). Although neither of these two PSR sequences matches any gene present in the *N. vitripennis* genome, PSR4317 shows some homology to a nuclear hormone receptor of several invertebrates and contains several putative functional domains, including a zinc finger C4 type DNA binding domain, RNA polymerase I specific transcription initiation factor RRN3 domain, and a transcription factor IIB zinc-binding domain. Interestingly, the highest homology covering the whole length of PSR4317 corresponds to a gene present in the endosymbiont *Candidatus cardinium*, which is known to produce reproductive manipulations, including cytoplasmic incompatibility (CI), parthenogenesis, and feminization, in certain arthropod host species (*30*). The observation that the majority of expressed transcripts encoded by PSR are most homologous to *T. sarcophagae* genes (Table S7, S8) is consistent with the previously proposed hypothesis that *Nasonia* PSR originated from the *Trichomalopsis* lineage by interspecific hybridization (*31*).

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

Using RT-qPCR (Real time quantitative PCR), we detected expression of these three genes in the male germ line, as well as in all tested embryonic and adult male somatic tissues, demonstrating that their expression is not regulated to a specific developmental time (Fig. S5). If any of these genes play a role in PSR's drive, their action on the paternally inherited chromatin could, in principle, occur either during sperm formation or instead in the egg's cytoplasm, immediately following fertilization. To address these possibilities, we used RNAi to transiently reduce transcript levels for each of these genes (PSR-4317, PSR-1539, PSR-tra) in the male germ line, and assessed for an effect on drive by measuring the sex ratio of F1 progeny (Fig. 2A, 2B). Knock-down of PSR4317 transcripts in the testes yielded a striking effect on F1 sex ratio: 46% of treated males produced broods that ranged between 10% and 90% F1 females (Fig. 2B). In contrast, RNAi targeting of the other two genes resulted in all-male broods similar to those from control (untreated) PSR+ wasps, despite the fact that their transcript levels were effectively reduced to levels comparable to RNAi-targeted PSR4317 transcript levels (Fig. 2A, 2B, Fig. S6). Maternal RNAi (*32*) targeting of PSR4317 in the egg's cytoplasm had no effect on F1 sex ratio (Fig. S7; see Methods). Thus, the effect of RNAi on sex ratio appears to be limited to the male germ line.

The production of F1 female progeny by RNAi-treated fathers could result from suppression of the genome eliminating activity, but it could also occur from destabilization and loss of PSR before the genome-eliminating activity has occurred. Several observations strongly argue against this latter possibility. First, individual sperm produced by RNAi-treated males contained a single copy of the PSR chromosome (Fig. 2C, Fig. S8), and PSR also was present in the nuclei of young fertilized embryos sired by these males (Fig. 2D, Fig. S9). Additionally, using PCR, we confirmed that the F1 adult females produced in our crosses were positive for multiple different PSR-specific sequences, suggesting that they inherited the B chromosome from their fathers (Fig. S7C, Fig. S10, Table S9). Together, these findings led us to speculate that the appearance of F1 female progeny may result from failure of paternal genome elimination when PSR4317 transcripts are knocked down, thus implicating PSR4317 as an active facilitator of PSR's drive.

To further test this possibility, we examined the mitotic behavior of the sperm- and egg-derived nuclei and their mitotic descendants in young embryos produced by RNAi-treated fathers. Normally, in embryos from control PSR+ fathers, the sperm-derived nuclear material fails to

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

resolve into individual chromosomes upon entry into the first mitosis, forming a mass of unresolved chromatin that remains distinct from the maternally derived nuclei ((*13*, *14*) Fig. 2D). This paternal chromatin mass (PCM) never segregates, eventually becoming lost within the embryo, as the egg-derived nuclei continue to divide (*13*, *14*). Remarkably, in nearly all of the embryos from PSR+ fathers RNAi-treated for PSR4317, there was no distinct PCM (Fig 2D, Fig. S9). Instead, in about 68% (26 out of 38) of these embryos the paternal chromatin formed bridges that spanned between dividing nuclei (Fig. 2D, Fig. S9). Such bridges were never observed in control PSR+ embryos (Fig. 2D). Additionally, some embryos from RNAi-treated fathers contained multiple nuclei that varied widely in size, suggesting that they were mosaics of nuclei with varying ploidy levels. In support of this idea, we observed one, two, and in some cases more rDNA (ribosomal DNA) foci per nucleus in these embryos (Figure 2D). We also examined two different histone post-translational modifications (PTMs), H3K9me3 (tri-methylation of histone 3 lysine 9) and H4K20me1 (methylation of histone 4 lysine 20), which under control (PSR+) conditions become abnormally distributed across the paternal chromatin in the presence of PSR (*11*). In young embryos from RNAi-treated fathers, however, both of these histone PTMs appeared more similar to patterns that are present in wild type (non-PSR) embryos (Fig. 2E). Taken together, these observations suggest that RNAi-targeting of PSR4317 alleviates the defective state of the sperm-derived chromatin caused by PSR, thereby allowing the sperm-derived chromatin to partially segregate with the egg-derived chromatin, giving rise to embryos that are mosaics of nuclei with differing amounts hereditary material derived from the two parents. Thus, we named this gene *haploidizer* because its expression is required for conversion of diploid embryos into haploids by causing paternal genome elimination.

We further speculated that the PSR+ females generated by RNAi-treated, PSR+ fathers arise from a portion of the mosaic embryos in which genome elimination is partially suppressed. Consistent with this hypothesis, using RT-PCR we were able to amplify both female- and male-specific transcript isoforms of the sex-determining genes, *transformer* (*tra*) and *doublesex* (*dsx*), in all tested PSR+ F1 females, but only male-specific isoforms in control PSR+ and wild type males (Fig. S10). In *N. vitripennis*, initiation of female sex determination is regulated epigenetically (*33*) and requires an unknown trans-acting factor termed *womanizer* (*wom*). This factor promotes female specific splicing of *transformer* (*traF*), whose presence, in turn, directs female-specific splicing of *doublesex* (*dsxF*), thereby promoting female development (*33*). It has

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

been hypothesized that *wom* is epigenetically silenced during oogenesis, and its absence in unfertilized eggs ensures male-specific development (*33*). We therefore propose that the PSR-induced alteration of the sperm-derived chromatin in control (PSR+) embryos is sufficient to block the production of *dsxF*, thereby facilitating male development. However, in fertilized embryos from RNAi-treated, PSR+ fathers, a less defective state of the paternal chromatin may allow some expression of *wom* and subsequently of *traF* and *dsxF* (Fig. 3). We note that F1 PSR+ females were unable to transmit PSR to their progeny, unlike their male siblings (Fig. 3). It is unlikely that the mosaic nature of these PSR+ females is the cause of this transmission failure because these individuals produce viable offspring, thus indicating normal transmission of the five essential chromosomes. Instead, this effect may reflect an intrinsic block to PSR transmission through the female germ line.

Although it is not yet known how *haploidizer*'s expression in the male germ line leads to alteration and eventual elimination of the sperm-derived chromatin, our study suggests an intriguing possibility: that *haploidizer* may be derived from a symbiont's gene that causes cytoplasmic incompatibility (CI) in its insect host. While speculative, this hypothesis is consistent with the fact that both symbiont-induced CI (*30*) and PSR's effect similarly involve disruption of the paternally inherited chromatin, and they occur at the same mitotic division.

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

Figures

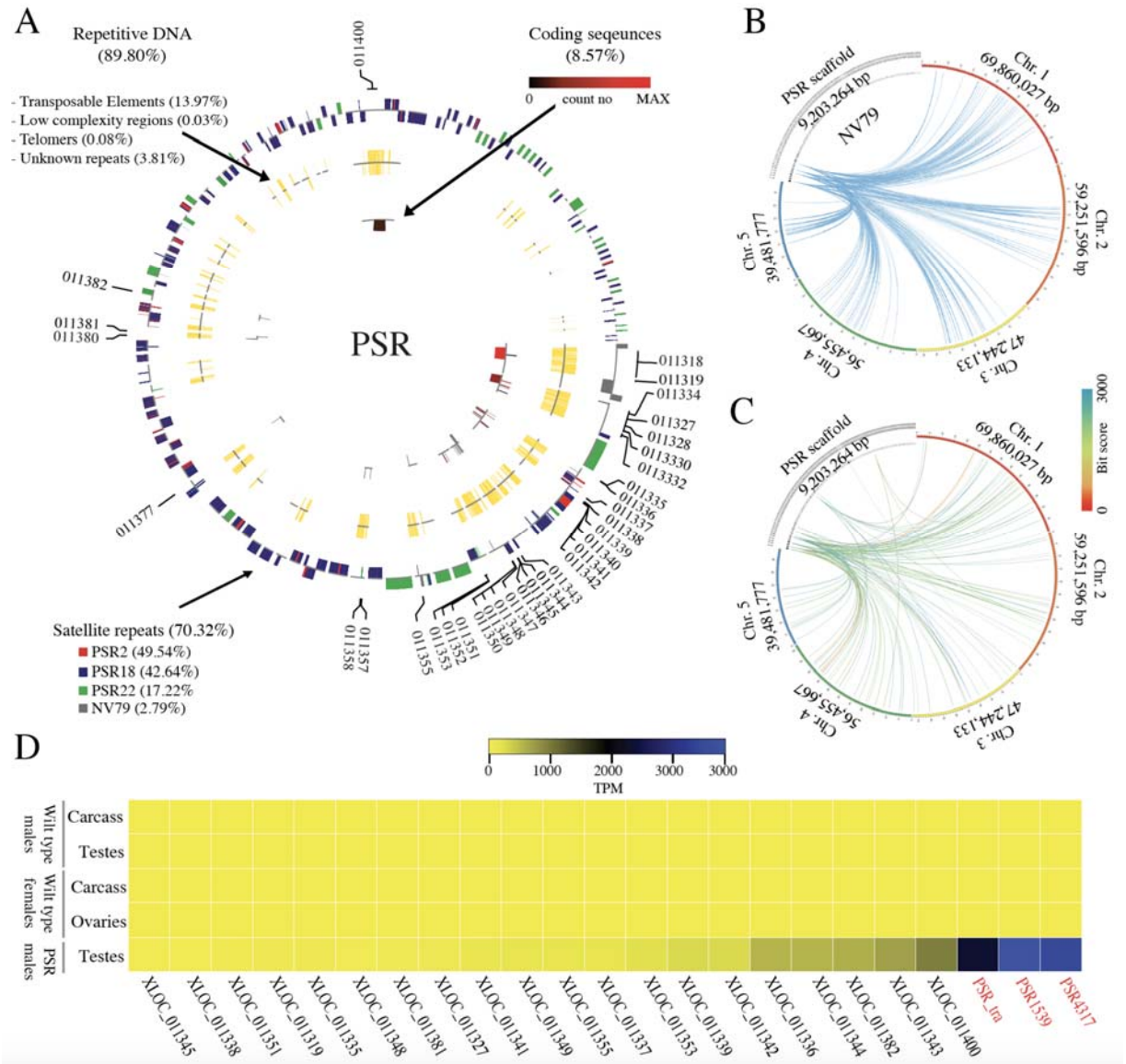

Fig. 1 Composition and origin of PSR chromosome

A) Circular visualization of the PSR chromosome scaffolds. The outermost circle represents repetitive satellites with outer bars representing the positive DNA strand and the inner bars representing those on the negative DNA strand. The colors represent the four major (70.32%) satellite families. The middle track represents other repetitive sequences (17.89%) including TE's, low complexity regions and telomeric sequences. The innermost track represents protein coding sequences (8.57%) and the color ranges from black to red for low to high expression, respectively. **B)** The relationship of NV79 repeats between PSR scaffolds and host chromosomes 1-5 indicating that PSR repeats are homologous to sequences found on all chromosomes. **C)** The relationship of PSR protein coding sequences and host chromosomes. The links are colored by blast bit scores from red (highest) and blue (lowest). Note: the PSR scaffolds are not placed in order and there is no correlation between the relationship in B and C despite the appearance of the graph. **D)** Heatmap of PSR gene with expression values higher than 10 TPM.

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

Expression ranges from yellow to blue from low to high expression. PSR genes that were disrupted with RNAi in this study are shown in red.

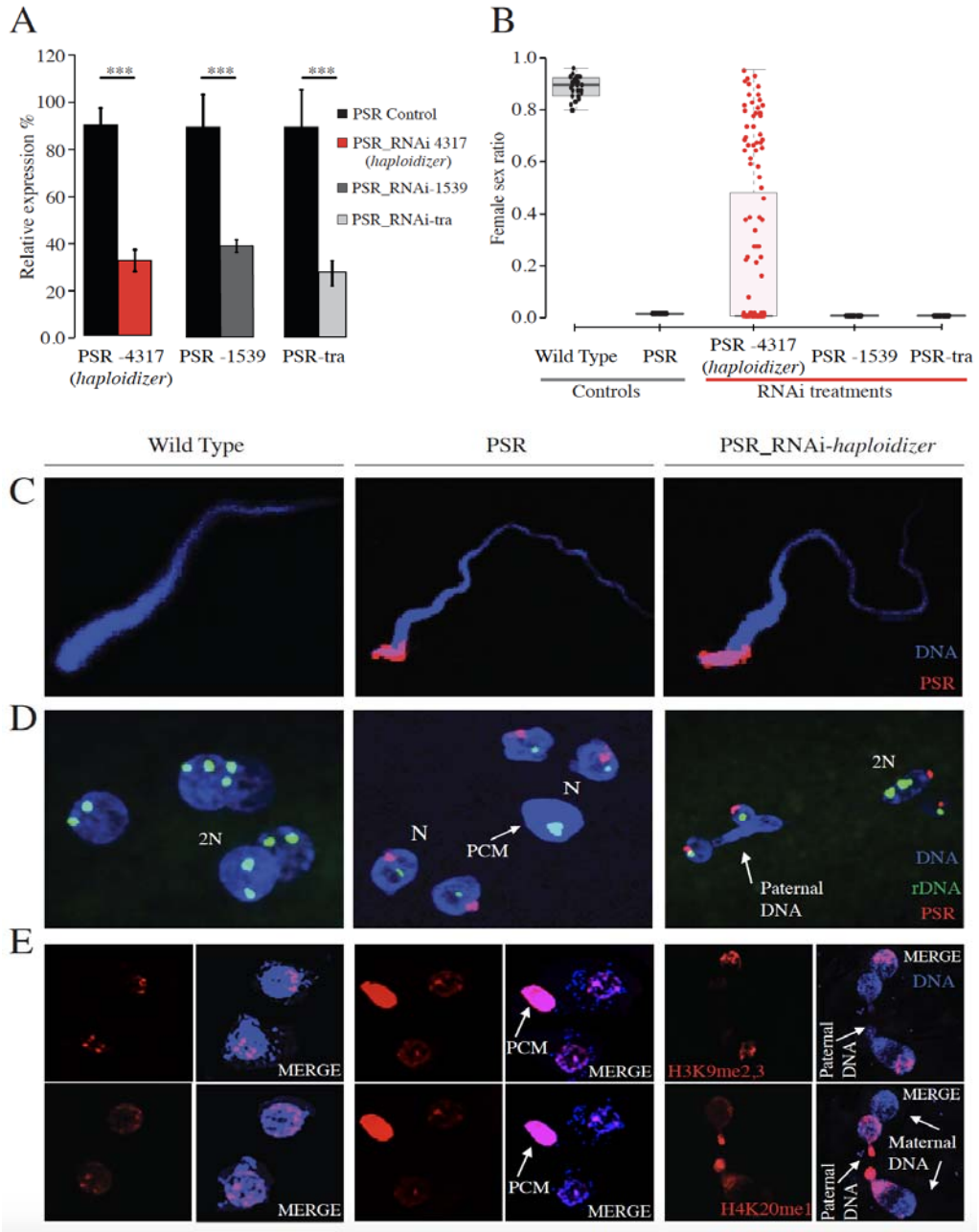

Fig. 2. Effects of *haploidizer* targeted by RNAi in early embryos

A) Efficiency of RNA interference measured by relative gene expression of the targeted genes. Asterisks indicate significant expression differences between control and RNAi-treated individuals (n=5, $P < 0.05$, one-way ANOVA with Tukey's multiple-comparisons test). **B)** Female F1 sex ratio was measured for controls and RNAi-treated wasps. The box plots depict the median (thick horizontal line within the box), the 25 and 75 percentiles (box margins) and the 1.5 interquartile range (thin horizontal line). **C)** Confocal imaging of sperm from wild type, PSR+ and *haploidizer* RNAi-treated PSR+ males. **D)** Embryos from wild type, PSR+ and *haploidizer* RNAi-treated PSR+ fathers, stained for with PSR (red) and rDNA (green). **E)** Embryos from wild type, PSR+ and *haploidizer* RNAi-

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

treated PSR+ fathers, stained for the histone marks, H3K9me2,3 (top panels, red) and H4K20me1 (bottom panels, also red). DNA is blue in all panels.. PCM= Paternal chromatin mass, 2N= diploid nuclei, N= Haploid nuclei.

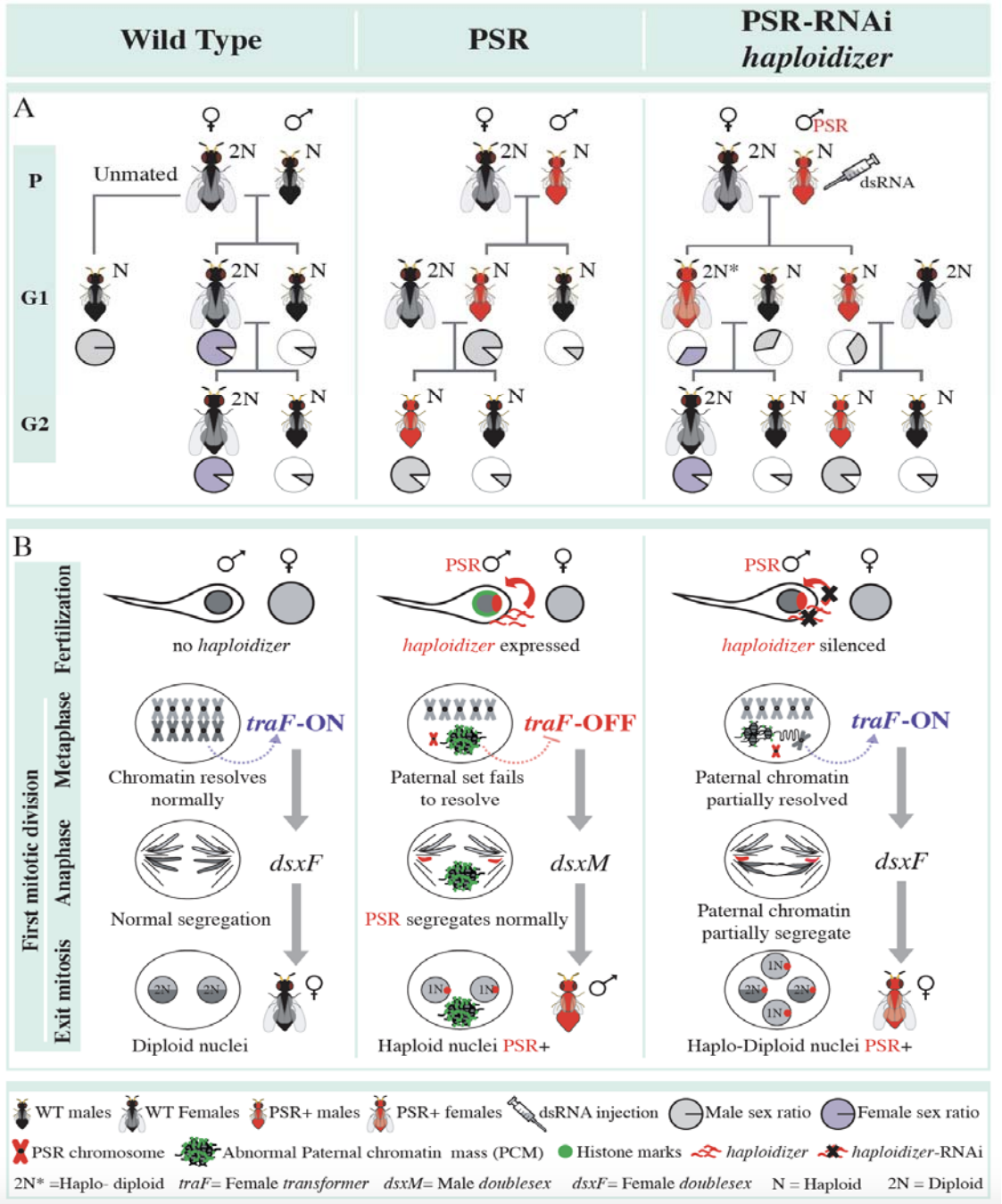

Fig. 3. Genetics and active model of PSR-induced genome elimination

A) The effects of RNAi knockdown of haploidizer on PSR transmission and sex ratio. The pie charts indicate male (grey) and female (purple) sex ratios **B)** Model for involvement of *haploidizer* in paternal genome elimination. Expression of PSR's *haploidizer* gene leads to failure of the paternal chromatin to resolve into chromosomes. This action inhibits the activation of female *transformer* (*traF*) and, thus, female development. As a result, these fertilized eggs develop into haploid male offspring that carry PSR. RNAi treatment of *haploidizer* alleviates the

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

chromatin defects, allowing partial expression of *traF* from the paternal set. This action re-establishes the female developmental pathway.

References

1. R. N. Jones, B-Chromosome Drive. *Am. Nat.* **137**, 430–442 (1991).

2. R. N. Jones, R. Neil Jones, B chromosomes in plants. *New Phytol.* **131**, 411–434 (1995).

3. J. P. M. Camacho, T. F. Sharbel, L. W. Beukeboom, B-chromosome evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci*. **355**, 163–178 (2000).

4. J. H. Werren, R. Stouthamer, PSR (paternal sex ratio) chromosomes: the ultimate selfish genetic elements. *Genetica*. **117**, 85–101 (2003).

5. G. D. D. Hurst, J. H. Werren, The role of selfish genetic elements in eukaryotic evolution. *Nature Reviews Genetics*. **2**, 597–606 (2001).

6. R. N. Jones, B-Chromosome Drive. *Am. Nat.* **137**, 430–442 (1991).

7. E. Dalla Benetta, O. S. Akbari, P. M. Ferree, Sequence expression of supernumerary B chromosomes: Function or fluff? *Genes* . **10** (2019).

8. M. Kimura, H. Kayano, The maintenance of super-numerary chromosomes in wild populations of *Lilium callosum* by preferential segregation. *Genetics*. **46**, 1699–1712 (1961).

9. K. M. Reed, Cytogenetic analysis of the paternal sex ratio chromosome of *Nasonia vitripennis*. *Genome*. **36**, 157–161 (1993).

10. K. M. Reed, J. H. Werren, Induction of paternal genome loss by the paternal-sex-ratio chromosome and cytoplasmic incompatibility bacteria (*Wolbachia*): A comparative study of early embryonic events. *Molecular Reproduction and Development*. **40**, 408–418 (1995).

11. J. C. Aldrich, A. Leibholz, M. S. Cheema, J. Ausió, P. M. Ferree, A “selfish” B chromosome induces genome elimination by disrupting the histone code in the jewel wasp *Nasonia vitripennis*. *Sci. Rep.* **7**, 42551 (2017).

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

12. J. C. Aldrich, P. M. Ferree, Genome silencing and elimination: Insights from a "Selfish" B chromosome. *Frontiers in Genetics*. **8**, 50 (2017).

13. M. M. Swim, K. E. Kaeding, P. M. Ferree, Impact of a selfish B chromosome on chromatin dynamics and nuclear organization in *Nasonia*. *J. Cell Sci.* **125**, 5241–5249 (2012).

14. J. H. Werren, U. Nur, D. Eickbush, An extrachromosomal factor causing loss of paternal chromosomes. *Nature*. **327**, 75–76 (1987).

15. J. H. Werren, L. W. Beukeboom, Population genetics of a parasitic chromosome: theoretical analysis of PSR in subdivided populations. *Am. Nat.* **142**, 224–241 (1993).

16. D. G. Eickbush, T. H. Eickbush, J. H. Werren, Molecular characterization of repetitive DNA sequences from a B chromosome. *Chromosoma*. **101**, 575–583 (1992).

17. L. W. Beukeboom, J. H. Werren, Deletion analysis of the selfish B chromosome, Paternal Sex Ratio (PSR), in the parasitic wasp *Nasonia vitripennis*. *Genetics*. **133**, 637–648 (1993).

18. O. S. Akbari, I. Antoshechkin, B. A. Hay, P. M. Ferree, Transcriptome profiling of *Nasonia vitripennis* testis reveals novel transcripts expressed from the selfish B chromosome, paternal sex ratio. *G3*. **3**, 1597–1605 (2013).

19. Y. Li, X. A. Jing, J. C. Aldrich, C. Clifford, J. Chen, O. S. Akbari, P. M. Ferree, Unique sequence organization and small RNA expression of a "selfish" B chromosome. *Chromosoma*. **126**, 753–768 (2017).

20. J. H. Werren, S. Richards, C. A. Desjardins, O. Niehuis, J. Gadau, J. K. Colbourne, Nasonia Genome Working Group, J. H. Werren, S. Richards, C. A. Desjardins, O. Niehuis, J. Gadau, J. K. Colbourne, L. W. Beukeboom, C. Desplan, C. G. Elsik, C. J. P. Grimmelikhuijzen, P. Kitts, J. A. Lynch, T. Murphy, D. C. S. G. Oliveira, C. D. Smith, L. van de Zande, K. C. Worley, E. M. Zdobnov, M. Aerts, S. Albert, V. H. Anaya, J. M. Anzola, A. R. Barchuk, S. K. Behura, A. N. Bera, M. R. Berenbaum, R. C. Bertossa, M. M. G. Bitondi, S. R. Bordenstein, P. Bork, E. Bornberg-Bauer, M. Brunain, G. Cazzamali, L. Chaboub, J. Chacko, D. Chavez, C. P. Childers, J.-H. Choi, M. E. Clark, C. Claudianos, R. A. Clinton, A. G. Cree, A. S. Cristino, P. M. Dang, A. C. Darby, D. C. de Graaf, B. Devreese, H. H.

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

Dinh, R. Edwards, N. Elango, E. Elhaik, O. Ermolaeva, J. D. Evans, S. Foret, G. R. Fowler, D. Gerlach, J. D. Gibson, D. G. Gilbert, D. Graur, S. Gründer, D. E. Hagen, Y. Han, F. Hauser, D. Hultmark, H. C. Hunter 4th, G. D. D. Hurst, S. N. Jhangian, H. Jiang, R. M. Johnson, A. K. Jones, T. Junier, T. Kadowaki, A. Kamping, Y. Kapustin, B. Kechavarzi, J. Kim, J. Kim, B. Kiryutin, T. Koevoets, C. L. Kovar, E. V. Kriventseva, R. Kucharski, H. Lee, S. L. Lee, K. Lees, L. R. Lewis, D. W. Loehlin, J. M. Logsdon Jr, J. A. Lopez, R. J. Lozado, D. Maglott, R. Maleszka, A. Mayampurath, D. J. Mazur, M. A. McClure, A. D. Moore, M. B. Morgan, J. Muller, M. C. Munoz-Torres, D. M. Muzny, L. V. Nazareth, S. Neupert, N. B. Nguyen, F. M. F. Nunes, J. G. Oakeshott, G. O. Okwuonu, B. A. Pannebakker, V. R. Pejaver, Z. Peng, S. C. Pratt, R. Predel, L.-L. Pu, H. Ranson, R. Raychoudhury, A. Rechtsteiner, J. T. Reese, J. G. Reid, M. Riddle, H. M. Robertson, J. Romero-Severson, M. Rosenberg, T. B. Sackton, D. B. Sattelle, H. Schlüns, T. Schmitt, M. Schneider, A. Schüler, A. M. Schurko, D. M. Shuker, Z. L. P. Simões, S. Sinha, Z. Smith, V. Solovyev, A. Souvorov, A. Springauf, E. Stafflinger, D. E. Stage, M. Stanke, Y. Tanaka, A. Telschow, C. Trent, S. Vattathil, E. C. Verhulst, L. Viljakainen, K. W. Wanner, R. M. Waterhouse, J. B. Whitfield, T. E. Wilkes, M. Williamson, J. H. Willis, F. Wolschin, S. Wyder, T. Yamada, S. V. Yi, C. N. Zecher, L. Zhang, R. A. Gibbs, Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. *Science*. **327**, 343–348 (2010).

21. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. **31**, 3210–3212 (2015).

22. A. B. Hall, Y. Qi, V. Timoshevskiy, M. V. Sharakhova, I. V. Sharakhov, Z. Tu, Six novel Y chromosome genes in *Anopheles mosquitoes* discovered by independently sequencing males and females. *BMC Genomics*. **14**, 273 (2013).

23. C. A. Desjardins, J. Gadau, J. A. Lopez, O. Niehuis, A. R. Avery, D. W. Loehlin, S. Richards, J. K. Colbourne, J. H. Werren, Fine-scale mapping of the *Nasonia* genome to chromosomes using a high-density genotyping microarray. *G3*. **3**, 205–215 (2013).

24. Smit, AFA, Hubley, R. RepeatModeler Open-1.0. 2008-2015

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

<http://www.repeatmasker.org>.

25. Smit, AFA, Hubley, R & Green, P. RepeatMasker Open-4.0. 2013-2015 <http://www.repeatmasker.org>.

26. M. Vítková, J. Král, W. Traut, J. Zrzavý, F. Marec, The evolutionary origin of insect telomeric repeats, (TTAGG) N. *Chromosome Research*. **13**, 145–156 (2005).

27. S. L. Hanlon, D. E. Miller, S. Eche, R. Scott Hawley, Origin, composition, and structure of the supernumerary B chromosome of *Drosophila melanogaster*. *Genetics*. **210**, 1197–1212 (2018).

28. A. Marques, S. Klemme, A. Houben, Evolution of plant B chromosome enriched sequences. *Genes*. **9**, 515 (2019).

29. I. Jetybayev, A. Bugrov, V. Dzuybenko, N. Rubtsov, B Chromosomes in grasshoppers: different origins and pathways to the modern Bs. *Genes*. **9**, 509 (2018).

30. A. Chigira, K. Miura, Detection of "*Candidatus Cardinium*" bacteria from the haploid host *Brevipalpus californicus* (Acari: Tenuipalpidae) and effect on the host. *Exp. Appl. Acarol*. **37**, 107–116 (2005).

31. B. F. McAllister, J. H. Werren, Hybrid origin of a B chromosome (PSR) in the parasitic wasp *Nasonia vitripennis*. *Chromosoma*. **106**, 243–253 (1997).

32. J. A. Lynch, C. Desplan, A method for parental RNA interference in the wasp *Nasonia vitripennis*. *Nat. Protoc.* **1**, 486–494 (2006).

33. E. C. Verhulst, J. A. Lynch, D. Bopp, L. W. Beukeboom, L. van de Zande, A new component of the *Nasonia* sex determining cascade is maternally silenced and regulates *transformer* expression. *PLoS One*. **8**, e63618 (2013).

34. S. Koren, B. P. Walenz, K. Berlin, J. R. Miller, N. H. Bergman, A. M. Phillippy, Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*. **27**, 722–736 (2017).

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

35. M. D. Cao, S. H. Nguyen, D. Ganesamoorthy, A. G. Elliott, M. A. Cooper, L. J. M. Coin, Scaffolding and completing genome assemblies in real-time with nanopore sequencing. *Nat. Commun.* **8**, 14515 (2017).

36. B. J. Walker, T. Abeel, T. Shea, M. Priest, A. Abouelliel, S. Sakthikumar, C. A. Cuomo, Q. Zeng, J. Wortman, S. K. Young, A. M. Earl, Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*. **9**, e112963 (2014).

37. H. Li, Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. **34**, 3094–3100 (2018).

38. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods*. **14**, 417–419 (2017).

39. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. **30**, 923–930 (2014).

40. E. Dalla Benetta, L. W. Beukeboom, L. van de Zande, Adaptive differences in circadian clock gene expression patterns and photoperiodic diapause induction in *Nasonia vitripennis*. *The American Naturalist*. **193**, 881–896 (2019).

41. J. Martin, T. Chong, P. M. Ferree, Male killing spiroplasma preferentially disrupts neural development in the *Drosophila melanogaster* Embryo. *PLoS ONE*. **8**, 79368 (2013).

42. A. M. Larracuente, P. M. Ferree, Simple method for Fluorescence DNA *in situ* hybridization to squashed chromosomes. *Journal of Visualized Experiments* ,**6**, 52288 (2015)

bioRxiv preprint first posted online Oct. 4, 2019; doi: http://dx.doi.org/10.1101/793273. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

Acknowledgments: We thank Vijaya Kumar for help with library preparations. Sequencing was performed at the Millard and Muriel Jacobs Genetics and Genomics Laboratory at the California Institute of Technology. **Funding:** This work was supported in part by funding from generous UCSD lab startup funds awarded to O.S.A. and NSF funds awarded to P.M.F. (NSF-1451839). **Author contributions:** O.S.A. and P.M.F. conceptualized the study. E.D.B., I.A., T.Y. and Q.M.N. performed the various genetics, molecular experiments and bioinformatic analysis in the study. All authors contributed to the writing, analyzed the data, and approved the final manuscript. **Competing interests:** All authors declare no significant competing financial, professional, or personal interests that might have influenced the performance or presentation of the work described. **Data and materials availability:** All Illumina, Pacbio and Oxford Nanopore sequencing data can be found at NCBI under accession number PRJNA575073 (https://dataview.ncbi.nlm.nih.gov/object/PRJNA575073?reviewer=uslhbl0iop93gab77lb31ts0n5). Previously published RNA-seq data (18) were also used in this study. They can be found under the accession number PRJNA212516 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA212516). The genome assembly submission ID is SUB6362372. All other data is available in the main text or the supplementary materials.

Supplementary Materials:

Materials and Methods

Figures S1-S11

Tables S1-S10

Data S1-S2

References (*34-42*)