Supplemental Methods

Sperm activation test

L4 males were picked out and placed into a new NGM plate for 72 h before dissection. Males were dissected and sperms released into a drop of sperm medium (SM) buffer (50 mM Hepes, 45 mM NaCl, 25 mM KCl, 1 mM $MgCl_2$, 5 mM $CaCl_2$, 10 mg/ml PVP, pH 7.0), then were maintained in chambers formed by mounting a 22×22-mm glass coverslip onto a glass slide with two parallel strips of two-sided sticky tape. To activate the sperm, various activators were diluted into the SM buffer and treated for 10 minutes before observation, Pronase (200 µg/ml, 10 minutes), zinc (1 mM, 15 minutes), monensin (100 nM, 15 minutes) and TEA (30 mM, 15minutes). After the treatments, sperms were observed under the Axio Imager M2 microscope (Carl Zeiss) with differential interference contrast (DIC).

MitoTracker staining and mating test

Mito-Tracker Red powders were diluted in DMSO to a final concentration of 3mM as stock solution, which was further diluted in M9 buffer to a final concentration of 0.1mM with 3.3% DMSO as working solution for worm socking. Worms were collected by washing plate and socked in MitoTracker working solution for 2.5 hours. GFP expressing males were picked under a stereo microscope. 10 sterile males were mated with 5 *C. nigoni* (JU1421) L4 females for 24 hours before taking image with a Leica SP5 confocal microscope.

Illumina long read generation, *C. nigoni* genome assembly and its annotation

Around 2.8-million long reads with a length of over 1.5 kb were produced as described (Li et al. 2015), with most of the reads sized around 10 kb, which covers about 20 times of *C. nigoni* genome. The reads were used to perform *de novo* assembly of *C. nigoni* genome using

Overlap Layout Consensus (OLC) based Celera Assembler (version 8.1) (Venter et al. 2001) using the same parameters as used for C. *elegans* long reads (Li et al. 2015).

To locate syntenic regions between *C. nigoni* and *C. briggsae genome*, the contigs assembled with long reads were aligned against *C. briggsae* genome ("cb4") using LAST (Kielbasa et al. 2011) to obtain one-to-one syntenic regions. The contigs were chained together to form *C. nigoni* pseudo chromosomes with reference to the *C. briggsae* syntenic order. The gap between two adjacent contigs was filled with one hundred "Ns". The remaining contigs, for which their one-to-one syntenic regions cannot be found in the *C. briggsae* genome, were merged as "un". The assembly termed as "cn1" is available from http://158.182.16.70:8080/cgi-bin/gb2/gbrowse_syn/cni/.

To annotate *C. nigoni* genome, a "bam" file was generated by mapping mRNA sequencing reads from JU1421 young adult males (3 libraries together) to the "cn1" genome. We used BAKER1 pipeline (Hoff et al. 2015) to perform gene prediction using "cn1" genome sequence and the "bam" file as inputs to produce genome annotation file in GFF format. To refine gene boundaries, we fed the *de novo* assembled transcripts and the GFF file into the Annotation Update Pipeline from PASA (Haas et al. 2008) to produce the final gene models as shown in http://158.182.16.70:8080/cgi-bin/gb2/gbrowse/cn1/. To calculate d_N/d_S ratio, after alignment of all protein sequence between *C. briggsae* and *C. nigini* using BLASTP, the 1:1 orthologous pairs were obtained as mutual best hits. Pair-wise codon alignment between 1:1 orthologous pairs was performed using MUSCLE (Edgar 2004), and the d_N/d_S ratios were calculated using KaKs_calculator (version 2.0) with MLPB (Modified LPB) method (Zhang et al. 2006).

Data analysis on mRNA-seq

After trimming the reads containing adaptors or with low quality (q score less than 20) using Trimmomatic (Bolger et al. 2014), the reads from all sources were mapped against *C.*

briggsae genome ("cb4") using CLC genomic workbench 8.0. The GFF annotation files for mRNA and gene annotation were obtained from Wormbase (WS250). To ensure the same mappability between *C. nigoni* and *C. briggsae* reads against *C. briggsae* ORFs, we relaxed the mismatch cutoff for *C. nigoni* and hybrid reads, which allows at least 85% similarity and 80% length while the default parameters, i.e., 90% similarity and 85% length, were used for *C. briggsae* reads. To find the misregulated/differentially expressed genes (DEGs) between the introgression strains and JU1421, we used edgeR package (Zhou et al. 2014) to calculate the fold change of RPKM and FDR.

We defined an expressed gene as one with at least 5 reads from each sample uniquely mapped onto its exonic regions in at least 3 out of all 12 samples as advised by the edgeR package (Zhou et al. 2014). We defined a significant change in expression as at least two-fold difference in average expression of 3 replicates with an FDR ≤ 0.01 between the hybrid and *C. nigoni* males. This approach could potentially compromise the mappability for *C. nigoni* genes that are diverged substantially from *C. briggsae*, but it allowed us to take advantage of the assembly and annotation of the *C. briggsae* genome. The mRNA reads of JU1421 and AF16 were first assembled *de novo* using Trinity with default parameters. The resulting transcriptome was refined by performing alignment against *C. briggsae* protein sequences using Blastp and Blastx. The JU1421 transcriptome was used for refining the *C. nigoni* gene models predicted by BAKER1 (Hoff et al. 2015) as well as for small RNA mapping described below.

Data analysis on microRNAs and piRNAs

Processing of small RNA sequencing was performed as described (Sarkies et al. 2015). Sequences identical to *C. briggsae* miRNAs (taken from miRbase) were identified using a custom PERL script, and the expression of these miRNAs compared between hybrid and *C. nigoni*. To check for miRNAs that might be missing in the parent but not the hybrid due to

sequence divergence we identified miRNAs where there was a clear discrepancy in read count (>4-fold difference) between the two. We then searched again through the small RNAs that had not been identified as perfect matches for miRNAs that had the same sequence within the seed region and >50% identity in the remaining region, using a second custom PERL script. The highest overall identity was used as a criterion to select the homologous miRNA, and this was then further tested by alignment to the *C. nigoni* genome. We then combined the reads from these small RNAs with the reads from the perfect matches to compare expression between *C. nigoni*, hybrid and *C. briggsae*. To investigate the precursor sequence and structure for miR-237, the full length hairpin was used to blast the *C. nigoni* contigs. The putative precursor region was extracted and we used RNAfold to predict secondary structure for this and the *C. briggsae* miRNA precursor. To examine the sequence divergence in the upstream region we used blastn to identify the homologous region in *C. briggsae* and align the two together.

To identify piRNAs, 21U-RNAs that did not map either exactly or with up to 1 mismatch to a *C. briggsae* miRNA were aligned to the *C. nigoni* genome assembly ("cn1") produced in this study using bowtie, allowing 0 mismatches and reporting only one alignment per sequence. We visually examined plots of the density of piRNA loci across the *C. nigoni* genome using histograms made in R, verifying that piRNAs were strongly enriched at the syntenic regions to the *C. briggsae* piRNA clusters described previously (Shi et al. 2013). We then prepared plots of the piRNA clusters to compare the number of unique sequences per million total unique sequences. To assess the difference in the overall piRNA read counts we used the Wilcoxon test (unpaired), which makes no assumption about the distribution involved.

To analyze 22G RNAs we aligned 22G RNAs to either TEs or transcripts from *C. nigoni* (See supplemental files at http://158.182.16.70:8080/share). We then assessed the total counts for 22G RNAs mapping to individual genes or TEs and this total read count was normalized to the total number of mapped 22G RNAs. To compare the 22G RNAs mapping to different

gene classes we first used annotations of 22G RNAs mapping to *C. briggsae* (Claycomb et al. 2009). This study reports CSR-1 targets in *C. briggsae* hermaphrodites by a direct immunoprecipitation approach, and also adds information about WAGO targets by homology. We supplemented these annotations with identification of CSR-1 targets in *C. elegans* males (Conine et al. 2013). We then used blast to identify the best matching *C. nigoni* transcript to the *C. briggsae* transcriptome, discarding genes that failed to map with an e-value of less than 10^{-4}, for which we could not assign a homologue with a better than 10^{-4} e-value. The significance of up or downregulation of 22G RNAs was assessed using the Wilcoxon test (paired between individual loci), which does not make any assumption of the underlying distribution.

Microscopy

DIC micrographs were acquired with hybrid detector of Leica SP5 confocal microscope equipped with water-immerse objective lens (63X).

Supplemental References

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*: btu170.

Claycomb JM, Batista PJ, Pang KM, Gu W, Vasale JJ, van Wolfswinkel JC, Chaves DA, Shirayama M, Mitani S, Ketting RF et al. 2009. The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. *Cell* 139(1): 123-134.

Conine CC, Moresco JJ, Gu W, Shirayama M, Conte D, Jr., Yates JR, 3rd, Mello CC. 2013. Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. elegans. *Cell* 155(7): 1532-1544.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* 32(5): 1792-1797.

Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. 2008. Automated eukaryotic gene structure annotation using EVidenceModeler and

the Program to Assemble Spliced Alignments. *Genome Biol* 9(1): R7.

Hoff KJ, Lange S, Lomsadze A, Borodovsky M, Stanke M. 2015. BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics*: btv661.

Kielbasa SM, Wan R, Sato K, Horton P, Frith MC. 2011. Adaptive seeds tame genomic sequence comparison. *Genome research* 21(3): 487-493.

Li R, Hsieh CL, Young A, Zhang Z, Ren X, Zhao Z. 2015. Illumina Synthetic Long Read Sequencing Allows Recovery of Missing Sequences even in the "Finished" C. elegans Genome. *Scientific reports* 5: 10814.

Sarkies P, Selkirk ME, Jones JT, Blok V, Boothby T, Goldstein B, Hanelt B, Ardila-Garcia A, Fast NM, Schiffer PM. 2015. Ancient and novel small RNA pathways compensate for the loss of piRNAs in multiple independent nematode lineages. *PLoS biology* 13(2): e1002061.

Shi Z, Montgomery TA, Qi Y, Ruvkun G. 2013. High-throughput sequencing reveals extraordinary fluidity of miRNA, piRNA, and siRNA pathways in nematodes. *Genome research* 23(3): 497-508.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA. 2001. The sequence of the human genome. *Science (New York, NY)* 291(5507): 1304-1351.

Zhang Z, Li J, Zhao X-Q, Wang J, Wong GK-S, Yu J. 2006. KaKs_Calculator: Calculating Ka and Ks Through Model Selection and Model Averaging. *Genomics, Proteomics & Bioinformatics* 4(4): 259-263.

Zhou X, Lindsay H, Robinson MD. 2014. Robustly detecting differential expression in RNA sequencing data using observation weights. *Nucleic acids research* 42(11): e91-e91.