

# <sup>1</sup>Selfing is the safest sex for *Caenorhabditis* <sup>2</sup>*tropicalis*

3

4Luke M. Noble<sup>1,2</sup>, John Yuen<sup>1</sup>, Lewis Stevens<sup>3</sup>, Nicolas Moya<sup>3</sup>, Riaad Persaud<sup>1</sup>, Marc Moscatelli<sup>1</sup>,  
5Jacqueline L. Jackson<sup>1</sup>, Gaotian Zhang<sup>3</sup>, Rojin Chitrakar<sup>6</sup>, L. Ryan Baugh<sup>6</sup>, Christian Braendle<sup>4</sup>, Erik C.  
6Andersen<sup>3</sup>, Hannah S. Seidel<sup>5</sup>, and Matthew V. Rockman<sup>1</sup>

7

81. Department of Biology and Center for Genomics & Systems Biology, New York University, New York,  
9NY, USA

102. Institute de Biologie, École Normale Supérieure, CNRS, Inserm, Paris, France.

113. Department of Molecular Biosciences, Northwestern University, Evanston, IL, USA

124. Institut de Biologie Valrose, Université Côte d'Azur, CNRS, Inserm, Nice, France.

135. Department of Biology, Eastern Michigan University, Ypsilanti, MI, USA.

146. Department of Biology, Duke University, Durham, North Carolina, USA.

## 15Abstract

16Mating systems have profound effects on genetic diversity and compatibility. The convergent evolution  
17of self-fertilization in three *Caenorhabditis* species provides a powerful lens to examine causes and  
18consequences of mating system transitions. Among the selfers, *C. tropicalis* is the least genetically  
19diverse and most afflicted by outbreeding depression. We generated a chromosomal-scale genome for  
20*C. tropicalis* and surveyed global diversity. Population structure is very strong, and islands of extreme  
21divergence punctuate a genomic background that is highly homogeneous around the globe.

22Outbreeding depression in the laboratory is caused largely by multiple Medea-like elements, genetically  
23consistent with maternal toxin/zygotic antidote systems. Loci with Medea activity harbor novel and  
24duplicated genes, and their activity is modified by mito-nuclear background. Segregating Medea  
25elements dramatically reduce fitness, and simulations show that selfing limits their spread. Frequent  
26selfing in *C. tropicalis* may therefore be a strategy to avoid Medea-mediated outbreeding depression.

## 27Introduction

28Sex and outcrossing are common, but costly, and taxa have repeatedly evolved mating systems to  
29avoid them. Such transitions set the advantages of selfing, such as reproductive assurance, against  
30long-term adaptability (Otto, 2009). Selfing has profound consequences for evolution due to changes in  
31effective recombination, homozygosity, and migration, leading to a net reduction in effective population  
32size. Mixed mating systems, combining some form of selfing with occasional outcrossing, are a  
33frequent compromise (Chelo et al., 2019; Cutter, 2019; Escobar et al., 2011; Goodwillie et al., 2005;  
34Igic & Kohn, 2006; Jarne & Auld, 2006).

35

36Variation in mating systems is especially familiar in plants, but has also been one of the longstanding  
37attractions of nematode biology (Nigon & Félix, 2017). Just within Rhabditidae, this aspect of life-history  
38now spans systems with separate males and females (gonochorism), males and self-fertile  
39hermaphrodites (andro dioecy; (Kanzaki et al., 2017; Mayer et al., 2007)), males, females, and  
40hermaphrodites (trioecy; (Chaudhuri et al., 2015; Kanzaki et al., 2017)), asexual reproduction where

41sperm does not contribute genetic material (parthenogenesis, gynogenesis; (Fradin et al., 2017;  
42Grosmaire et al., 2019)), and alternating generations of hermaphroditism and dioecy (heterogony;  
43(Kiontke, 2005)). Within the *Caenorhabditis* genus, the androdioecious system of males and self-  
44fertilizing hermaphrodites has evolved three times independently (Ellis, 2017). Hermaphrodites are  
45morphologically female, but during larval development they generate and store sperm for use as adults.  
46Hermaphrodites cannot mate with one another, and in the absence of males all reproduction is by self  
47fertilization.

48

49*C. tropicalis* was identified by Marie-Anne Félix in 2008 (Kiontke et al., 2011), and investigations of its  
50reproductive biology have shed light on the mechanistic basis of transitions to selfing (Wei et al., 2014;  
51Zhao et al., 2018). However, a comprehensive reference genome for the species is lacking and  
52relatively little is known of its biology and ecology. Global sampling indicates a more restricted range  
53than that of the other selfers (Félix, 2020), and the single study of *C. tropicalis* population genetics and  
54reproductive compatibility found extremely low levels of genetic diversity at a handful of loci (Gimond et  
55al., 2013). Crosses among, and sometimes within, locales revealed outbreeding depression, a result  
56common to the selfers (Baird & Stonesifer, 2012; Dolgin et al., 2007; Ross et al., 2011) but in stark  
57contrast to gonochoristic species, where estimated diversity is often orders of magnitude higher and  
58inbreeding depression can be severe (Barrière et al., 2009; Dolgin et al., 2007; Gimond et al., 2013).  
59Outbreeding depression was particularly acute in *C. tropicalis*, which exhibits frequent embryonic  
60lethality and developmentally abnormal F<sub>2</sub> progeny among certain hybrid crosses (Gimond et al., 2013).  
61Male mating ability was also found to be generally poor, though highly variable, which together with low  
62genetic diversity suggests an especially high rate of selfing.

63

64Theoretical explanations for the evolution of selfing consider a balance between benefits and costs.  
65Classical models start with the doubling of reproductive rate achieved by selfers (through the  
66elimination of males in the case of *Caenorhabditis*) and note that selfing will evolve when the reduction  
67in reproduction due to inbreeding depression is less than one half (Cutter, 2019; Goodwillie et al., 2005;  
68Lande & Schemske, 1985; Lively & Lloyd, 1990). Subsequent work has added a variety of factors on  
69both sides of the balance. Selfing has benefits for reproductive assurance (Baker, 1955; Theologidis et  
70al., 2014), allowing a single individual to colonize a new environment, and it allows for rapid adaptation  
71when traits are jointly determined by maternal and zygotic genotypes (Drown & Wade, 2014). On the  
72other hand, selfing slows adaptation to new environments or pathogens by reducing recombination and  
73genetic diversity, while increasing the load of weakly deleterious mutations (Cutter, 2019; Kamran-  
74Disfani & Agrawal, 2014; Morran et al., 2009, 2011). Many features of the selective environment (e.g.,  
75stabilizing vs. directional) and the genetic architecture of fitness (e.g., recessive deleterious variation  
76vs. overdominance as the cause of inbreeding depression) shape the balance of factors, and vary  
77among species and populations (Goodwillie et al., 2005). The ledger of relevant factors runs long, but  
78as we show, it may still be incomplete.

79

80To study the causes and effects of mating system evolution in *Caenorhabditis*, we assembled a  
81chromosome-scale genome for *C. tropicalis*, oriented by linkage data from recombinant inbred lines.  
82We show with short-read mapping against this reference that the population structure of a global  
83sample of isolates is very strong, which is an expected side-effect of selfing. We also find extreme  
84heterogeneity in the distribution of genetic diversity across the genome. Finally, we investigate the  
85causes of strong outbreeding depression, another expected side-effect of selfing, in a cross between  
86divergent isolates. Widespread outbreeding depression observed in the three species of selfing  
87*Caenorhabditis* has been interpreted as evidence for the well understood process of Dobzhansky-

88Muller epistasis. Here, we show that it is due largely to a different process in *C. tropicalis*: maternal-  
89effect haplotypes that kill offspring that do not inherit them. Haplotypes that behave in this way are  
90known as Medea elements (Beeman et al., 1992; Beeman & Friesen, 1999), and they represent a form  
91of post-zygotic gene drive (Tom A. R. Price et al., 2020; Wade & Beeman, 1994). *C. tropicalis* Medea  
92elements impose severe fitness costs on heterozygous mothers. We hypothesize that Medea elements  
93in *C. tropicalis* select for a high selfing rate in this species, by adding another factor to the balance on  
94the side of selfing. In other words, frequent selfing may be a consequence of outbreeding depression  
95as much as its cause.

## 96Results

### 97Heritable variation in outcrossing rates

98To better understand variation in outcrossing in *C. tropicalis*, we tested a global sample of five isolates  
99for their propensity or ability to outcross. For each of these strains, from Hawaii, Panama, French  
100Guiana, Cape Verde, and Réunion Island, we assayed the probability that an individual would produce  
101cross progeny when paired with a single individual of the other sex. Mating success was scored as a  
102binary trait, with crosses scored as successfully mated if multiple male offspring were observed on the  
103plate. This scoring system allowed for the possibility of single male offspring to be produced by X  
104nondisjunction in hermaphrodites, which typically occurs at a low rate (~1% or lower, see below). All  
105factorial crosses produced progeny, but strains varied significantly in their propensity to outcross, both  
106as males and as hermaphrodites (Figure 1A). We also observed interaction effects, where the average  
107crossing probability of male and hermaphrodite strains was not predictive of the success of the strains  
108in combination. At one extreme, the pairing of JU1373 hermaphrodites from Réunion with JU1630  
109males from Cape Verde, yielded no cross progeny from 22 trials, while at the other extreme, Hawaiian  
110QG131 hermaphrodites and South American NIC58 males yielded cross progeny in each of 34 trials.  
111Male crossing probability was much more variable than that of hermaphrodites, consistent with relaxed  
112selection on XO males (residual deviance of 179.2 vs. 85.2, null deviance 232.6, binomial linear model)  
113(Cutter, 2019; Jalinsky et al., 2020; Noble et al., 2015; Palopoli et al., 2008; van der Kooi & Schwander,  
1142014). Same-strain pairings were indistinguishable from inter-strain pairings ( $p=0.88$ , likelihood ratio  
115test (LRT) of binomial linear models). This analysis shows that wild isolates of *C. tropicalis*, and males  
116in particular, vary greatly in their propensity or ability to outcross.  
117

118Given extensive variation in outcrossing and the independence of male and hermaphrodite components  
119of this trait, we asked whether strains could maintain males over time, or whether hermaphrodite selfing  
120would drive them from populations (Steward & Phillips, 2002). We founded single-strain populations  
121with three hermaphrodites and five males, allowed them to expand, and then serially transferred their  
122descendants for 10 generations under standard lab conditions. This experiment tests additional  
123components of fitness, such as fecundity, in a continuously expanding population. At the end, two  
124strains, JU1630 and JU1373, had lost males and were reproducing solely by selfing; these are the  
125strains with the lowest cross success in our single-worm pairings (Figure 1B). The other strains retained  
126males at frequencies of 20-40%. This dichotomy resembles that seen among *C. elegans* strains, where  
127some strains become exclusive selfers, while others can maintain high rates of outcrossing in simple  
128environments (Steward & Phillips, 2002; Teotónio et al., 2006; Teotonio et al., 2012). We chose the  
129strain with the most-outcrossing males, NIC58 from French Guiana, and the strain with the least-  
130outcrossing hermaphrodites, JU1373 from Réunion Island, as focal strains for an investigation of the  
131genetics and population biology of *C. tropicalis*.

133Figure 1. (A) Outcrossing probability in reciprocal crosses. Mating success was scored as a binary trait in 22-34  
 134trials (biological replicates) per cross. Marginal means with bootstrap 99% confidence intervals are shown. (B)  
 135After 10 generations of passaging at large population size, strains vary in their male frequency (mean and  
 136standard error of three biological replicates). R: Réunion Island, CV: Cape Verde, FG: French Guiana, H: Hawaii,  
 137Pa: Panama. Data are in Figure 1 - source data 1 and 2.

### 138A chromosomal genome for NIC58

139To generate a reference genome for *C. tropicalis* we used deep PacBio long-read sequencing of  
 140NIC58, and then applied genetic linkage data from recombinant inbred lines (RILs) to assess and orient  
 141multiple assemblies. RILs were derived from a cross between a JU1373 hermaphrodite and a NIC58  
 142male, by selfing F<sub>2</sub> hermaphrodites from a single F<sub>1</sub> for 10 generations. We genotyped lines by shotgun  
 143sequencing, called diallelic single nucleotide variants (SNVs), and inferred parental ancestry by Hidden  
 144Markov Model, using data for 119 RILs for genetic map estimation. During RIL construction, 12.2% of  
 145lines did not survive, consistent with outbreeding depression. The genotypes of the surviving RILs  
 146revealed strong transmission ratio distortion favoring JU1373 alleles at two loci on chromosomes III and  
 147V (Figure 5), which we return to later.

148

149Using the RIL recombination data to evaluate assemblies, we selected one of five fully concordant with  
 150the genetic data. We then closed gaps of estimated 0 cM distance, using junction-spanning sequences  
 151in other assemblies, followed by local long-read mapping. Five gaps of greater than 0 cM remain, which  
 152will require longer reads to resolve. The resulting nuclear genome assembly comprised 81.3 Mb in 15  
 153sequences. The X chromosome assembled as a single contig, and all other chromosomes were  
 154oriented genetically into pseudochromosomes. We assembled a 13,935 bp mitochondrial genome from  
 155short-reads, followed by circular extension with long-reads. To better assess genetic variation between  
 156the RIL founder strains, we also assembled draft nuclear (81 Mb span, 4.2 Mb NG50) and  
 157mitochondrial (13,911 bp) genomes for JU1373 from similar read data (see Methods; we do not attempt  
 158to bring this assembly to pseudochromosomes here). We annotated nuclear genomes using mixed-  
 159stage short-read RNAseq data, calling 21,210 protein-coding genes for NIC58 and 20,829 for JU1373,  
 160and we annotated the mitochondrial genomes by homology. Thus, this pipeline provided a high-quality  
 161assembly for NIC58 and a highly contiguous draft genome for JU1373.

### 162Surveying genetic diversity worldwide

163*C. tropicalis* is widely distributed within 25° of the Equator and absent outside this region (Félix, 2020).  
 164To begin a global survey of genetic diversity and population structure, we sequenced an additional 22  
 165isolates that broadly represent the species' global range with short reads (Figure 3A). The collection  
 166spans Africa, Asia, and South and Central America, but large equatorial regions, notably in Central  
 167Africa and Southeast Asia, are not yet represented. Our sample included 16 American isolates (eight  
 168from the Caribbean, eight from Central and South America), four from East Asia, three from Africa and  
 169one from the Central Pacific (Figure 3 - source data 1). We called variants against the NIC58 reference  
 170genome, hard-filtered to 794,676 diallelic SNVs on the nuclear genome (genotype set 1,  
 171Supplementary File 4; see Methods) and selected 397,515 sites with fully homozygous calls and no  
 172missing data (genotype set 2, Supplementary File 5) for exploration of population structure. We called  
 173mitochondrial variants separately, and filtered similarly, retaining 166 (of 197 hard-filtered) SNVs.  
 174

175 Previously Gimond et al. (2013) sequenced 5.9 kb across nine nuclear, protein-coding loci in 54  
176 isolates (mostly from French Guiana in South America, but including African isolates from Cape Verde  
177 and Réunion Island, and our Pacific isolate from Hawaii) and found nine SNVs, equating to a per-site  
178 Watterson's  $\theta$  around 0.00034. Though not directly comparable, the genome-wide estimate of  
179 nucleotide diversity provided here is around three times higher (genotype set 1, median value across  
180 20 kb windows = 0.00097), with mitochondrial diversity higher again as expected (0.0038). These  
181 values likely underestimate species-wide variation because of short-read mapping bias – we adjust for  
182 missing data, but missing data may not, in fact, be missing from genomes. Indeed, we found rampant  
183 missingness in our data; up to 1.3% of the NIC58-alignable fraction of the genome lacks aligned reads  
184 among any single isolate, and 7.8% of the alignable genome lacks reads in at least one (considering  
185 only those with >25x mapped and paired reads, for which missing data and sequencing depth are  
186 uncorrelated,  $r = 0.13$ ,  $p = 0.65$ ). These results establish that *C. tropicalis* shows very low genetic  
187 diversity genome-wide, but also suggest this homogeneity is broken by many regions sufficiently  
188 divergent that reference-based mapping fails.

### 189 ***C. tropicalis* genetic diversity is highly heterogeneous along chromosomes**

190 Selfing has complex effects on population dynamics and genome evolution. A general expectation is a  
191 reduction in effective population size  $N_e$  proportional to the frequency of selfing, by a factor of up to two,  
192 and the strength of background selection, by potentially much more (Charlesworth, 2012; Nordborg &  
193 Donnelly, 1997). The strength of selection acting on genetic variation is proportional to the product of  
194  $N_e$  and the selection coefficient  $s$ . Selfing is therefore expected to both lower genetic diversity, making  
195 evolution more reliant on new mutations, and raise the threshold below which mutations are effectively  
196 neutral. We found that the distribution of SNV diversity along all chromosomes is extremely  
197 heterogeneous in *C. tropicalis*. Recombination is relatively homogenous within the large recombination  
198 rate domains, at least in *C. elegans* (Bernstein & Rockman, 2016; Kaur & Rockman, 2014), and is  
199 therefore not expected to generate heterogeneity in the effects of linked selection at these small scales.  
200 Background diversity in *C. tropicalis* (median  $\theta_w$ ) is the lowest among the three selfers, and the median  
201 number of SNV differences between isolates ( $\pi$ ) in 10 kb windows is just 3.2 on chromosome centers,  
202 and less than double that on arms (genotype set 2). Variance around the background is almost eight  
203 times that of *C. elegans* and more than 100 times that of *C. briggsae* (data in Figure 2A). We used  
204 kernel density smoothing of the binned distribution of  $\theta_w$  to partition the genome into segments of very  
205 high diversity (the long right tail of divergent outlier regions) and segments with background levels of  
206 diversity (e.g., see Figure 2A). Heterogeneity is often highly localized: at 10 kb scale, 141 divergent  
207 peaks fall to background within 30 kb or less (see Methods), and divergent regions cover just over 14%  
208 of the NIC58 genome in sum. Genetic diversity in *C. tropicalis* is thus typified by a near-invariant  
209 background, suggesting very recent global shared ancestry, punctuated by regions of high divergence,  
210 consistent with a possible role for balancing selection.

211  
212 Finally, to gain a view of *C. tropicalis* genetic diversity less subject to reference-mapping bias, we  
213 aligned the draft JU1373 genome against NIC58, calling variants and assessing copy number variation  
214 from alignment depth (H. Li, 2018). From 78.86 Mb of aligned bases (69.18 at single copy), we saw a  
215 37% increase in SNVs over reference-based mapping, and a sum of 1.23 Mb in insertion-deletion  
216 variation including 388 variants of length greater than 1 kb. SNV divergence in 10 kb windows  
217 commonly exceeded 10% on the arms (Figure 2B), and total divergence (the sum of variant length  
218 differences relative to NIC58) exceeded 30% in windows on every chromosome. Reference-based SNV  
219 calling thus dramatically underestimated the true levels of genetic diversity at divergent loci, which were  
220 comparable to current estimates for gonochoristic species and to analogous patterns recently described

221in *C. elegans* and *C. briggsae* (Lee et al., 2020). In the face of such high divergence, at single base and  
222multigene-scale, long-read genomes and variant graph genome representation may be required to  
223more fully describe species-wide variation (Garrison et al., 2018). Deeper population genetic data may  
224also allow inference of divergent foci within these loci.

225

226Figure 2. (A) Nucleotide diversity across chromosomes based on 24 strains for *C. tropicalis*, 35 strains for *C.*  
227*briggsae*, and 330 isotypes for *C. elegans* (Watterson's  $\theta$  in non-overlapping 20 kb windows  $\times 100$ , normalised  
228across species and chromosomes). The denominator in Watterson's estimator uses the mean number of strains  
229with non-missing calls per window rounded to the nearest integer. 12 outliers for *C. elegans* are outside the  
230plotted range. Differences in heterogeneity across species are apparent from the marginal density plots, and from  
231dispersion around the locally-weighted polynomial (LOESS) fit to the data in black. Levels of variation at loci in *C.*  
232*tropicalis* centers approach those of arms for chromosomes II, IV, V, and X. Arm recombination rate domains are  
233shaded, and regions on the left arms of chromosome II and the X are magnified below. Here, triangles are local  
234peaks called at 10 kb scale by segmenting divergent regions (red) from background (blue) at the threshold  
235indicated by a grey line (see Methods), the y-axis is as in the main plot, and the the x-axis is unnormalised  
236physical distance (Mb). Data are in Figure 2 - source data 1. (B) Genetic diversity between JU1373 and NIC58  
237from genome alignment, shown as SNV differences (1 - % identity) in 10 kb non-overlapping windows. Data are in  
238Figure 2 - source data 2.

### 239***C. tropicalis* shows strong continental population structure with local heterogeneity**

240To examine population structure worldwide, we decomposed NIC58-reference-based genetic  
241relatedness of the 24 isolates into its principal components. We observed strong structure, with the top  
242axis differentiating three African isolates from all others and accounting for almost 75% of nuclear  
243genetic relatedness (Figure 3B). The close clustering of the three African lines, isolated across a  
244transect spanning more than 9000 km from the Atlantic island of Cape Verde, off the Westernmost  
245coast of continental Africa, to Réunion Island in the Indian Ocean, is remarkable given the large  
246geographic distances separating these locales, though it is consistent with the occurrence of globally  
247distributed haplotypes in *C. elegans* (Lee et al., 2020). The second principal component differentiates  
248Western Pacific samples (Hawaii, Taiwan) from all others, and the third principal component largely  
249differentiates two of four Taiwanese samples from Hawaii. These three dimensions account for 89% of  
250the genetic variance, which is of similar magnitude to the variance explained by the first three principal  
251components in *C. briggsae* (based on 449,216 SNVs with no missing data among 34 lines).

252

253The genome-wide view masks heterogeneity at the chromosome level. Notably, chromosome IV shows  
254much more complex patterns of relatedness, and both chromosome V and the mitochondrial genome  
255provide evidence of recent admixture, with strain QG834 from Panama clustering with strains from East  
256Asia and the Pacific. Structuring of the X is particularly extreme, with essentially two haplotypes present  
257in our sample, African and non-African, and this split accounts for 98% of the genetic variance (Figure 3  
258- figure supplement 1 and 2).

259

260While the structure revealed by this analysis explains most of the genetic variance, we also observe  
261variance within these populations at divergent regions. A minority of loci are divergent both within and  
262across populations (Figure 3C). We also find a handful of highly divergent regions that vary only within  
263populations. For example, loci on chromosomes IV and V vary among seven isolates from a single  
264collection made over one month from the small island nation of Dominica. The structure of *C. tropicalis*  
265populations therefore combines strong, global differentiation, as seen between tropical and temperate  
266*C. briggsae* clades; widespread homogeneity, as seen in *C. elegans* outside the Pacific; and diversity at

267the very local scale, common to all three selfing species (Andersen et al., 2012; Barrière & Félix, 2007;  
268Félix et al., 2013; Haber et al., 2005; Sivasundar & Hey, 2005).

269

270Figure 3. (A) The distribution of 24 isolates (numbers per locale are shown in the legend; Data in Figure 3 -  
271source data 1), colored by groupings in (B), where principal component analysis of nuclear genomic similarity  
272identifies largely discrete populations. (C) Genetic diversity is found mostly, but not entirely, within populations.  
273For three populations with at least two lines (<99% SNV identity), within population diversity ( $\pi$ ) is plotted for each  
274population against between population diversity ( $D_{xy}$ ; Nei & Li (1979); 10 kb scale). Data in B-C are based on  
275Supplementary File 4.

276

277Figure 3 - figure supplement 1. Chromosome and mitochondrial population structure. Point size for each  
278chromosome/genome panel scales with the percentage of variance explained (PVE) by the first two principal  
279components of SNV genetic relatedness. Data are based on Figure 3 - source data 1.

280

281

282Figure 3 - figure supplement 2. The average number of SNV differences among all pairwise comparisons within  
283( $\pi$ ; upper, with the global value for all pooled samples plotted in blue across each panel) and between (Nei and  
284Li's  $D_{xy}$  (1979); lower) populations (10 kb non-overlapping windows, adjusted for the mean fraction of missing  
285data per window). We use the major population groupings defined by genome-wide PCA in Figure 3B, where  
286multiple isolates are present (13 from Central and South America, 3 from Africa and 2 from East Asia after filtering  
287to <99% identity). Data are based on Figure 3 - source data 1.

## 288Quantitative genetics of outcrossing

289As a first step toward understanding the genetic basis of variation in outcrossing, we scored  
290hermaphrodites from the NIC58 x JU1373 RILs for their probability of producing cross offspring in  
291matings with NIC58 males. As expected from their differences in outcrossing rate in single worm and  
292bulk passaging assays (Figure 1), we observed considerable variation among lines, including  
293transgressive segregation (Figure 4A). Linkage mapping detected a significant effect of a locus on the  
294center of the X chromosome (Figure 4B), which explained close to 15% of the variance in  
295hermaphrodite outcrossing probability. Although the difference in equilibrium male frequency between  
296JU1373 and NIC58 is likely mediated by their different outcrossing rates, we also observed differences  
297in the rate of spontaneous male production due to X nondisjunction during hermaphrodite meiosis. Self  
298progeny of NIC58 hermaphrodites were 0.8% male (21/2580), versus 0.06% in JU1373 (2/3088); given  
299a characteristic brood size of 100-150, these numbers imply that most NIC58 self broods include a  
300male, and most JU1373 self broods do not. These two strains thus differ heritably in male crossing  
301ability, hermaphrodite crossing ability, equilibrium sex ratio, and spontaneous male production rate,  
302providing multiple paths for the evolution of outcrossing rate.

303

304Figure 4. (A) RILs vary in their hermaphrodite crossing probability. Means and 95% bootstrap confidence intervals  
305from binary trials are shown for RILs and their parents. Data are in Figure 4 - source data 1. (B) Quantitative trait  
306locus mapping for hermaphrodite crossing probability (genome-wide 0.05 significance threshold from 1000  
307phenotype permutations shown in grey, n = 118 RILs). Data are based on Supplementary File 1.

### 308RIL transmission ratio distortion and excess heterozygosity

309The RIL genotypes displayed strong transmission ratio distortion in two genomic regions: the left arm of  
310chromosome III and the right arm of chromosome V (Figure 5A). Both were strongly skewed toward  
311JU1373 homozygotes, which reached a frequency of 97% on chromosome III and 78% on chromosome  
312V. The chromosome V locus also showed an excess of heterozygotes (16% of RILs) compared to the  
313neutral expectation after 10 generations of selfing ( $1/2^{10} = 0.1\%$ ). RILs that retained heterozygosity on  
314chromosome V also showed an enrichment of JU1373 genotypes on chromosome I (18.2-20.7 cM),  
315which itself showed mild distortion in favor of JU1373 (Fisher's exact test, p < 0.001, Figure 5 - source  
316data 1). These data indicate that selection during RIL construction strongly favored JU1373 alleles on  
317chromosomes III and V, with complex selection at the locus on chromosome V favoring heterozygotes  
318over JU1373 homozygotes under some conditions.

319

320Figure 5. Two genomic regions show strong transmission ratio distortion. (A) RIL genotype frequencies and peaks  
321of transmission ratio distortion (arrows). Shaded areas are 1 LOD drop intervals and peak point estimates.  
322Genome-wide data are based on Supplementary File 1, and multilocus transmission ratio distortion genotype  
323tables are in Figure 5 - source data 1. (B) Fold coverage and SNV divergence of JU1373 relative to NIC58. Fold  
324coverage is in 5 kb windows, divergence is 1-identity in 100 bp windows. Data are based on Figure 2 - source  
325data 2.

### 326Transmission ratio distortion is not due to simple mitochondrial-nuclear 327incompatibilities

328Strong selection during RIL construction is consistent with previous reports of extensive outbreeding  
329depression (Gimond et al., 2013). Yet simple genetic incompatibility between two nuclear loci is not  
330expected to favor one parental allele to the exclusion of the other. Exclusion of one parental allele can  
331occur, however, if an allele from the male parent (NIC58) is incompatible with the mitochondrial  
332genome of the hermaphrodite parent (JU1373). Under this scenario, RILs homozygous for the male  
333parent allele at loci showing transmission ratio distortion should be sub-viable or sub-fertile. We  
334examined RILs of such genotypes and found that their growth characteristics were superficially normal,  
335with 93.1-98.3% (n = 151-679) of embryos developing into adults with parental developmental timing  
336(Figure 7A). This finding indicates that the cause of transmission ratio distortion was more complex  
337than a simple mitochondrial-nuclear incompatibility between these NIC58 alleles and JU1373  
338mitochondria.

### 339Transmission ratio distortion and excess heterozygosity are caused by Medea loci

340An alternative explanation for transmission ratio distortion among RILs is that the distorted loci  
341independently exhibit post-zygotic killing dynamics, similar to those seen in *C. elegans* at the *zeel-*  
342*1/peel-1* and *pha-1/sup-35* loci (Ben-David et al., 2017; Seidel et al., 2008). At these loci, a maternal- or  
343paternal-effect locus loads a toxin into eggs or sperm that poisons zygotic development; subsequent  
344zygotic expression from the same locus provides an antidote (Ben-David et al., 2017; Seidel et al.,

3452011). Independent of the precise mechanism, loci with this pattern of parental-by-zygotic genetics of  
346lethality are classified as Medea-type elements, named for their maternal-effect dominant embryonic  
347arrest (Beeman et al., 1992). We show below that the *C. tropicalis* distorted loci fit this mode of  
348inheritance, and broaden the Medea classification here to encompass post-embryonic arrest. A Medea  
349model predicts that the JU1373 genome is homozygous for two independent Medea haplotypes, on  
350chromosomes III and V, each encoding a hypothetical toxin-antidote pair. (In our analysis of this model  
351we describe it as a toxin-antidote system for convenience, though the underlying mechanism of  
352maternal-effect dominant lethality and zygotic-effect dominant rescue may be different.) Under this  
353model (Figure 5 - figure supplement 1), all F<sub>2</sub> progeny from a NIC58 x JU1373 cross will be exposed to  
354toxins, but only some will inherit antidotes; animals not inheriting both antidotes will suffer the effects,  
355manifesting as embryonic or larval arrest, sterility, or some other phenotype that would have prevented  
356them from contributing to the RILs. The proportion of F<sub>2</sub>s showing such phenotypes is expected to be  
357around 44% (7/16), assuming that arrest by each toxin is fully penetrant and rescuable by a single copy  
358of the corresponding antidote. Given the presence among RILs of rare NIC58 homozygotes at each of  
359the two transmission ratio distortion loci, penetrance must be less than 100%, and so the proportion of  
360affected F<sub>2</sub> progeny should be less than 44%.

361

362Figure 5 - figure supplement 1. Models of Medea-mediated gene drive and expected F<sub>2</sub> arrest proportions. Each  
363Medea factor is composed of two genes, a maternally expressed toxin and a zygotically expressed antidote. (A)  
364Two JU1373 Medeas, on chromosomes III and V. This model is not the best fit for observed genotype frequencies  
365among wild-type F<sub>2</sub> animals. (B) A JU1373 Medea on III and antagonistic Medeas (JU1373 and NIC58) on V. The  
366NIC58 Medea is assumed to be 50% penetrant, but true penetrance likely depends on genetic background. This  
367model is the best fit for observed genotype frequencies among wild-type F<sub>2</sub> animals in a mito-JU1373  
368background. (C) Inactive JU1373 Medea on III and antagonistic Medeas (JU1373 and NIC58) on V. This model is  
369the best fit for observed genotype frequencies among wild-type F<sub>2</sub> animals in a mito-NIC58 background.

370

371To test whether F<sub>2</sub> populations from a NIC58 x JU1373 cross showed phenotypes consistent with two  
372Medea-like loci, we made reciprocal crosses, allowed F<sub>1</sub> hermaphrodites to self-fertilize, and followed  
373F<sub>2</sub> progeny from embryo to adulthood. We observed that far fewer F<sub>2</sub> embryos developed into adults  
374within the normal developmental time than expected from the two-Medea model (41-45%, n = 329-  
3751283, versus 98-99%, n = 1046-1093, for parental strains). Terminal phenotypes among abnormal F<sub>2</sub>  
376animals included failure to hatch (9%), early larval arrest (39-41%), late larval arrest (5-9%), and  
377abnormally small, thin adults (5%, n = 329-381); these phenotype frequencies did not differ according  
378to the direction of the cross that generated the F<sub>1</sub> worm (Fisher's exact test, p = 0.39). These data show  
379that F<sub>2</sub> populations experienced widespread developmental arrest, at proportions higher than expected  
380under a model of two Medea loci. One interpretation is that the total arrest proportion among F<sub>2</sub>  
381individuals reflects the effects of two Medea loci (<44% of F<sub>2</sub> animals), plus additional background  
382incompatibilities (>12-15% of F<sub>2</sub> animals).

383

384The two-Medea model predicts that developmental arrest among F<sub>2</sub> animals will preferentially affect  
385those homozygous for non-Medea (NIC58) alleles. To test this prediction, we repeated reciprocal  
386NIC58 x JU1373 crosses and genotyped F<sub>2</sub> progeny (including arrested or dead larvae) at markers  
387tightly linked to the transmission ratio distortion peaks on chromosomes III and V. We observed that  
388alleles at both loci were transmitted to progeny in Mendelian proportions, showing that meiosis and  
389fertilization are not affected by the loci and that distortion is due to postzygotic lethality. Certain  
390genotypes were associated with developmental arrest, but these differed depending on the direction of  
391the cross (i.e., mitochondrial genotype). For the chromosome III locus, NIC58 homozygotes underwent

392complete developmental arrest in the mito-JU1373 cross but very little developmental arrest, compared  
393to other genotypes, in the mito-NIC58 cross (Figure 6A). For the chromosome V locus, NIC58  
394homozygotes experienced highly penetrant developmental arrest in both crosses (Figure 6A).  
395Chromosome V JU1373 homozygotes also experienced developmental arrest in both crosses, but with  
396a lower penetrance, especially in the mito-JU1373 cross (Figure 6A). This pattern supports a model of  
397two Medea loci, but implies that the loci interact with mitochondrial genotype: for the locus on  
398chromosome III, the Medea allele (JU1373) is active in its own mitochondrial background but inactive or  
399very weakly active in the opposite mitochondrial background; for the locus on chromosome V, both  
400alleles (JU1373 and NIC58) act as Medea elements and are effectively antagonistic – JU1373 acts  
401strongly in both mitochondrial backgrounds, whereas NIC58 acts strongly in its own mitochondrial  
402background and more weakly in the opposite mitochondrial background (Figure 6C). Importantly, this  
403finding of antagonistic haplotypes at the chromosome V locus provides a simple explanation for the  
404retention of heterozygosity at this locus among the RILs, since both classes of homozygous segregants  
405from a hermaphrodite heterozygous for the chromosome V locus arrest (Figure 5 - figure supplement  
4061).

#### 407**Antagonistic Medea activity does not reflect mitochondrial-nuclear incompatibility**

408The reciprocal crosses described above showed that developmental arrest of chromosome V JU1373  
409homozygotes was more penetrant in a mito-NIC58 background than in a mito-JU1373 background. One  
410model for this pattern is that any animal homozygous for the JU1373 chromosome V locus in the mito-  
411NIC58 background suffers a fitness cost due to a mito-nuclear incompatibility, independent of Medea  
412effects. Alternatively, there is no such intrinsic incompatibility, and instead the mitochondrial  
413background modifies the penetrance of the NIC58 chromosome V Medea, with effects manifest in the  
414offspring of chromosome V heterozygotes. The fitness of JU1373 V homozygotes in a mito-NIC58  
415background should be reduced under the mitochondrial-nuclear incompatibility model, independent of  
416the genotypes of their parents. We therefore attempted to isolate this genotype by crossing males of  
417RIL QG2514 to NIC58, producing an F<sub>2</sub> population segregating at the chromosome V locus but fixed for  
418the NIC58 haplotype at the chromosome III locus (Figure 6B). We recovered 10 F<sub>2</sub> animals  
419homozygous for the JU1373 V allele in a mito-NIC58 background. These animals produced progeny  
420that were superficially wild-type, with typical brood sizes and developmental rates. This finding excludes  
421simple mitochondrial-nuclear incompatibility as a contributor to developmental arrest of chromosome V  
422JU1373 homozygotes. We conclude that differences in arrest according to mitochondrial genotype  
423reflect greater activity of the NIC58 Medea on chromosome V in a mito-NIC58 background than in a  
424mito-JU1373 background. We also conclude that activity of the NIC58 Medea in the mito-JU1373  
425background is likely dependent on additional nuclear background factors, thus accounting for  
426heterozygosity in the RILs being preferentially retained in certain genetic backgrounds (e.g., in animals  
427with JU1373 alleles on the left arm of chromosome I). Finally, we note that similar conclusions  
428(mitochondrial x parental x zygotic interactions, not simple mitochondrial x zygotic incompatibility) can  
429be inferred for the chromosome III locus, given the wild-type development observed in the subset of  
430RILs homozygous for the NIC58 chromosome III locus in the mito-JU1373 background (Figure 7A).  
431Thus, lethality in *C. tropicalis* is not caused by simple mitochondrial x zygotic incompatibilities, although  
432two of the three Medea elements are more active in the mitochondrial background of their parental  
433strain (Figure 6C).

#### 434**Medea loci act independently**

435Our model predicts that the Medea loci act independently and are thus genetically separable. To test  
436this prediction, we examined RILs with opposite parental genotypes at the two loci. Each RIL was  
437crossed to an appropriate parental strain to generate F<sub>2</sub> animals segregating opposite alleles at one  
438Medea locus but fixed at the other, which we scored for development. To provide a control for (non-  
439Medea) background incompatibilities, we repeated this analysis for F<sub>2</sub> populations fixed for JU1373  
440alleles at both Medea loci. These controls were generated by crossing 10 RILs carrying JU1373 alleles  
441at both Medea loci to a JU1373-derived strain carrying a Dumpy mutation, which allows us to  
442distinguish cross progeny from self progeny, a constant issue for JU1373 with its low rate of crossing by  
443hermaphrodites (Figure 1). We observed that crosses segregating only the chromosome III Medea in a  
444mito-JU1373 background showed median rates of normal development consistent with Medea activity  
445at a single locus (68-75%, n = 218-240); crosses in a mito-NIC58 background showed no such activity  
446(Figure 7A). The arrested progeny in the mito-JU1373 background largely consisted of chromosome III  
447NIC58 homozygotes, as evidenced by these animals being severely depleted among wild-type progeny  
448(26:64:2 JJ:JN>NN genotypes), while the wild-type progeny of the reciprocal cross carried chromosome  
449III genotypes at their expected Mendelian proportions (30:57:31). These data show that the  
450chromosome III Medea is active in the absence of segregation at the chromosome V locus, and confirm  
451that the chromosome III Medea requires a mito-JU1373 genetic background. The lethality associated  
452with chromosome III therefore represents a three-way interaction between parental nuclear genotype,  
453zygotic nuclear genotype, and a mitochondrial locus.

454

455Crosses segregating only at the chromosome V locus showed median rates of normal development  
456that were similar across mitochondrial backgrounds and consistent with antagonistic Medea activity at a  
457single locus (54-61%, n = 203-321, in a mito-JU1373 background; 52-53%, n = 313-345, in a mito-  
458NIC58 background; Figure 7A). Arrested progeny largely consisted of NIC58 homozygotes and, to a  
459lesser extent, JU1373 homozygotes, as evidenced by these genotypes being depleted among wild-type  
460progeny (Figure 7B). This result shows that antagonistic Medea haplotypes at the chromosome V locus  
461act in the absence of variation at the chromosome III locus. For comparison, crosses segregating no  
462Medeas showed median rates of normal development that were usually higher than 75%, as expected  
463for crosses lacking drive activity (Figure 7A); nonetheless, rates of normal development in these  
464crosses were highly variable and sometimes low, with medians ranging from 65-100% (n = 163-213),  
465suggesting a contribution from non-Medea background incompatibilities. We conclude that the  
466chromosome III and V Medea loci act independently of one other, and that additional background  
467incompatibilities are widespread and diffuse in the NIC58 x JU1373 cross.

#### 468**Medea loci act via maternal effect**

469The loci we identified in *C. tropicalis* produce transmission ratio distortion via an interaction between  
470parent and offspring genotypes. To distinguish between maternal and paternal effects, we scored rates  
471of normal development among progeny in which Medea alleles segregated from mothers or fathers. We  
472used a combination of sperm depletion and visible markers to prevent self progeny from contaminating  
473counts of cross progeny. For a maternal effect (and no paternal effect), 50% of cross-progeny lacking  
474the Medea were expected to undergo developmental arrest when the Medea was present in the  
475mother, assuming full penetrance, but develop normally when the Medea was present in the father. In  
476the case of a paternal effect (and no maternal effect), the opposite pattern was expected. Crosses  
477testing the JU1373 and NIC58 Medeas showed that all acted via maternal, and not paternal, effect  
478(Figure 7B). The overall model is presented in Figure 5 - figure supplement 1 and Figure 7D.

479

480Figure 6. Medea genetics. (A) Percentage of F<sub>2</sub> progeny from reciprocal NIC58 x JU1373 crosses showing wild-  
481type development. Genotypes reflect markers tightly linked to the peaks of transmission ratio distortion on  
482chromosomes III and V. ? = genotyping failure. (B) Genotype and phenotype frequencies among F<sub>2</sub> progeny from  
483reciprocal crosses between NIC58 and RIL QG2514. Only wild-type F<sub>2</sub> progeny were genotyped. Data are in  
484Figure 6 - source data 1.

485

486Figure 7. Medea loci act independently and by maternal effect. (A-B) Percentages of F<sub>2</sub> or backcross progeny that  
487reached adulthood within 72 hours of egg laying. Strains beginning with "QG" are RILs. JU1373d is a Dumpy  
488mutant. NIC58r and QG4249r express a red fluorescent transgene (see Methods. Data in Figure 7 - source data  
4891). (A) Crosses testing whether Medea loci act independently. Each point is a cross plate, progeny of a single F<sub>1</sub>  
490hermaphrodite, with a median of 34 worms scored per plate. (B) Crosses testing whether Medea loci act via  
491maternal or paternal effect. Maternal- and paternal-effect expectations are under a model that either a maternal-  
492or paternal-effect toxin causes fully penetrant developmental arrest for progeny not inheriting the Medea  
493haplotype. (C) Interpretation of maternal- and paternal-effect crosses. (D) Schematic of Medea activity.  
494Percentages are estimates for the proportion of animals undergoing developmental arrest, compared to  
495heterozygous siblings, derived by comparing observed genotype frequencies among wild-type F<sub>2</sub> progeny to  
496Mendelian expectations. This method avoids bias introduced by genotyping failures being more common among  
497arrested versus wild-type animals. Reciprocal crosses used to estimate arrest proportions were NIC58 x JU1373,  
498NIC58 x RIL QG2479 (not shown), and NIC58r x RIL QG2514. NIC58r expresses a red fluorescent transgene  
499(see Methods).

## 500Incompatibilities are associated with duplicated and novel genes

501To better understand the genetic basis for Medea activity, we mapped the JU1373 elements using RILs  
502with recombination breakpoints near the peaks of transmission ratio distortion. RILs were crossed to  
503derivatives of parental strains, and Medea activity was assessed by scoring the rate of normal  
504development among F<sub>2</sub> progeny. This analysis mapped Medeas to intervals of around 33 kb on  
505chromosome III and 69 kb on chromosome V (Figure 7 - figure supplement 1) that overlapped the  
506peaks of transmission ratio distortion (Figure 5). The chromosome III interval encompasses a locus of  
507locally elevated sequence divergence between NIC58 and JU1373, in an otherwise well-conserved  
508region. A clear candidate is an insertion of sequence in JU1373 that includes seven predicted genes  
509(Figure 8A): six are tandem duplications of neighboring genes in NIC58; the seventh is a duplication of  
510a gene located 0.68 Mb to the right in both NIC58 and JU1373. Five of the seven genes have no  
511detectable protein or nucleotide homology outside *C. tropicalis*. The remaining two are homologous to  
512*C. elegans* genes *F44E2.8* and *F40F8.11*, which share NADAR and YbiA-like superfamily protein  
513domains. In addition to these seven genes, JU1373 also carries an eighth gene inserted within the  
514original copy of the duplicated sequence (Figure 8A, arrowhead). This gene is novel and has no  
515detectable homology to any gene in NIC58 or in any other species. Thus, the JU1373 Medea on  
516chromosome III contains a total of eight additional genes compared to NIC58: one unique to JU1373,  
517and five with no homology outside *C. tropicalis*.

518

519The chromosome V locus lies in a region of high divergence between NIC58 and JU1373, extending  
520well beyond the mapped interval (Figure 5B). This region is home to a number of dynamically evolving  
521gene families (Figure 8B). A major structural difference between JU1373 and NIC58 haplotypes is an  
522expansion of divergent F-box-domain-encoding genes, expanded from three homologs in NIC58 to 13  
523in JU1373. Immediately flanking this expansion in JU1373 is a duplication of a gene located 0.6 Mb  
524away (and present in both NIC58 and JU1373), which encodes a homolog of the checkpoint kinase  
525*chk-2*. Adjacent to the *chk-2* homolog is a tandem duplication of a nuclear gene encoding a  
526mitochondrial ubiquinol-cytochrome c oxidoreductase complex protein, perhaps accounting for the

527interaction of Medea activity with mitochondrial genotype. Additional differences between haplotypes  
528include a gene encoding a small, novel protein in JU1373 and a novel gene in NIC58. Thus, the  
529JU1373 Medea on chromosome V resides within a larger genomic window of elevated sequence  
530divergence, containing an expansion of F-box genes, a novel gene, as well as a duplicated gene  
531encoding a mitochondrial protein. The inferred antagonistic Medea in NIC58 may also involve novel  
532protein-coding potential.

533

534We also examined the mitochondrial genomes of JU1373 and NIC58, and found that the core functional  
535complement is superficially identical; all 12 protein-coding, 2 ribosomal RNA and 22 tRNA genes are  
536called as present and intact in both (Bernt et al., 2013; Lemire, 2005; R. Li et al., 2018). There are,  
537however, several differences of unknown significance, including 12 missense variants in six of the core  
538protein coding genes, the presence in NIC58 of a small non-coding region and a potentially novel (lowly  
539expressed) gene encoding a 2-pass transmembrane protein, and differential expression at two  
540additional loci (Figure 8 - figure supplement 1). These are additional candidate variants for the  
541observed nuclear-cytoplasmic interaction.

542

543Figure 7 - figure supplement 1. Mapping Medea loci using RILs. (A-B) Percentages of F<sub>2</sub> progeny that reached  
544adulthood within 72 hours of egg laying. Strains beginning with "QG" are RILs. JU1373d is a Dumpy mutant.  
545NIC58r expresses a red fluorescent transgene (see Methods; Data in Figure 7 - source data 1). (A) Crosses to  
546determine whether RILs with recombination breakpoints near Medea loci exhibit Medea activity in crosses to  
547JU1373d and NIC58r. ? = unknown RIL genotypes at the Medea loci, to be inferred from the results of the  
548crosses. Expectations are based on crosses shown in (B). Some of the data here are duplicated from Figure 7 for  
549ease of comparison. (C) RIL genotypes and intervals to which Medeas were mapped. Peak point intervals are  
550regions of maximum RIL distortion  $\chi^2$  statistics from Figure 5A.

551

552Figure 8. Genomic intervals surrounding Medea loci on chromosome III (A) and V (B). Windows span  
553transmission ratio distortion 1 LOD drop intervals. Rectangles are predicted genes, lines connect homologs.  
554Colors indicate the union of homologs within the interval. Homology relationships to genes outside the intervals  
555are not shown. Data are based on Supplementary File 2 and Supplementary File 6.

556

557Figure 8 - figure supplement 1. Mitochondrial genomes. Annotations are shown for two homology-based methods,  
558Mitos (Bernt et al., 2013) and Prokka (Seemann, 2014), over expression from short-read RNAseq data (per base  
559read depth). RO: potential replication origin, including the large D-loop non-coding region at 13.5 kb. Some  
560obvious differences of unknown significance are highlighted with black arrowheads along the three x-axes  
561including (top to bottom, left to right): the presence of a small, low-scoring RO in NIC58 called by Mitos; the  
562presence of a novel open reading frame, called by Prokka, overlapping the D-loop in NIC58 and predicted to  
563encode a 122-amino acid transmembrane protein; and three regions of clear differential expression overlapping a  
56416S ribosomal RNA fragment, a tRNA and RO cluster, and the D-loop region. Data are based on Supplementary  
565File 2.

## 566**Incompatibilities consistent with Medea activity are widespread among wild isolates**

567To examine the distribution of putative Medea activity among wild strains, we crossed 14 isolates to a  
568Dumpy derivative of JU1373 (JU1373d), to NIC58, or to both, and scored development among the F<sub>2</sub>.  
569Crosses were classified as having putative Medea activity if rates of normal development were below  
570~75% (the expectation for segregation of a single Medea). We observed that crosses to JU1373d  
571showed putative Medea activity for 12 of 13 isolates (Figure 9A), and the strength of activity was

572associated with geographic origin, and haplotype at the chromosome III and V Medea loci (which are  
573themselves associated due to population structure). The two African isolates showed the least activity  
574and had haplotypes similar to JU1373 at both Medea loci; five American isolates showed consistently  
575higher activity and carried haplotypes dissimilar to JU1373 at both Medea loci; isolates from other areas  
576were more variable, but activity was lowest for isolates carrying haplotypes similar to JU1373 at the  
577chromosome V locus (NIC535 and NIC773) or somewhat similar to JU1373 at both loci (QG131; Figure  
5789B). Nonetheless, the correlation between Medea activity, geographic origin, and haplotype was  
579imperfect (e.g., JU1630 versus JU3170), and some isolates showed median rates of normal  
580development consistent with segregation of three or more Medeas (NIC517 and QG834, 20-28%, n =  
581257-276). This pattern suggests that the JU1373 Medeas may be fixed within Africa but polymorphic or  
582absent elsewhere, and that some of the non-African isolates may contribute additional Medeas of their  
583own.

584

585Among crosses to NIC58, putative Medea activity was observed in six of nine crosses (Figure 9A). Five  
586of these crosses showed activity that was moderate in strength and largely overlapping, despite these  
587isolates having different haplotype combinations at the chromosome III and chromosome V Medea loci.  
588The sixth cross (JU1630) showed much higher activity, consistent with this isolate alone having  
589JU1373-like haplotypes at both Medea loci. The three remaining crosses showed no drive activity and  
590had haplotypes similar to NIC58 at both loci; we conservatively interpret these crosses as inconclusive  
591because we cannot be certain of paternity for F<sub>1</sub> animals (crosses to NIC58 did not include a marker to  
592distinguish cross- from self-progeny). Nonetheless, these data as a whole indicate that putative Medea  
593activity is widespread: Medeas sometimes segregate within a geographic region, and some crosses  
594may segregate additional elements to those identified in the NIC58 x JU1373 cross.

595

596Figure 9. Wild isolate phenotypes and haplotypes. (A) Percent of F<sub>2</sub> progeny reaching adulthood within the normal  
597developmental time (~72 hrs), for crosses between wild isolates and a Dumpy derivative of JU1373 (JU1373d) or  
598NIC58. #, inconclusive because we cannot be certain that F<sub>1</sub> animals were cross progeny. nd, not determined.  
599Putative Medea activity is inferred for crosses in which the median percent of F<sub>2</sub> animals reaching adulthood was  
600less than ~75%. Each plotted point is a plate (2-16 per cross, median 6), with a median 90 animals scored per  
601plate. Data is in Figure 9 - source data 1. (B) Wild isolate haplotypes at the Medea loci on chromosomes III and V.  
602Heterozygous calls likely reflect duplication and divergence. Data are based on Supplementary File 4.

### 603Medea dynamics in partial selfers

604Populations segregating Medea loci like those we discovered on chromosomes III and V are expected  
605to evolve suppressors to alleviate the reduced fitness of animals whose progeny are killed. A general  
606resistance mechanism for suppressing this lethality is selfing, which reduces the prevalence of  
607heterozygotes and hence of animals expressing Medea-associated lethality. To better understand how  
608the mating system of *C. tropicalis* influences the spread of these elements, we simulated evolution  
609under different levels of selfing and outcrossing. These simulations included incomplete penetrance as  
610well as idiosyncratic features of *Caenorhabditis* androdioecy: hermaphrodites cannot mate with one  
611another and males can arise spontaneously by nondisjunction of the X chromosome (see Methods).  
612Selfing rate was a fixed parameter in this system, and we leave a fuller exploration of the  
613coevolutionary dynamics of selfing and intergenerational incompatibilities to future work. We found that  
614in a large population, high rates of selfing can dramatically slow the rate of spread of a Medea element,  
615and indeed can reduce its efficacy such that it fails to increase in frequency during 100 generations of  
616simulated evolution (Figure 10A). In small populations, elements introduced at 5% frequency are rapidly  
617fixed under random mating but lost to drift under high rates of selfing (Figure 10B). These simulations

618also reveal that once an element is at high frequency, selfing actually hastens fixation of the Medea  
619element (compare partial selfing at  $S = 0.25$  to obligate outcrossing at  $S = 0$ ). This occurs because  
620once an element is at high frequency, heterozygotes that self will always expose their progeny to the  
621element's killing effects, but heterozygotes that outcross will usually mate with an antidote-carrying  
622male, slowing the fixation of the element. These results show that partial selfing slows or prevents the  
623spread of Medea elements at low frequency but hastens their fixation at high frequency.

624

625We next considered antagonistic Medeas, under the simplistic scenario of perfect linkage and equal  
626penetrance. We found that intermediate levels of selfing cause frequency-dependent selection against  
627the rarer haplotype (Figure 10C, lower panel, and Figure 10 - figure supplement 2A-C). Surprisingly,  
628allele frequencies of the antagonistic Medeas drift as though neutral under both obligate selfing and  
629obligate outcrossing, albeit for different reasons. With obligate selfing, there are simply no  
630heterozygotes and the alleles are literally neutral. Under obligate outcrossing, overdominance  
631generated in the offspring generation by the killing of both homozygote classes is offset by  
632underdominance in the parental generation, because the costs of these deaths are borne by  
633heterozygote mothers. The net effect is that the alleles exhibit drift-like dynamics despite the enormous  
634selective cost. Effective neutrality in this scenario depends on the equal penetrance of antagonistic  
635haplotypes in the model. When we model unequal penetrances, similar to those we observe at the  
636chromosome V locus (0.95 and 0.6), we find that a resident strong Medea will prevent the spread of a  
637weaker invader, as expected. When a strong Medea invades a population with a weaker resident  
638Medea, the strong haplotype spreads if outcrossing rates are high but is eliminated under moderate to  
639high rates of selfing (Figure 10 - figure supplement 2D, E). Overall, selfing reduces Medea load both by  
640decreasing heterozygote frequency and by inducing strong positive frequency-dependent selection that  
641prevents antagonistic alleles from co-occurring.

642

643Finally, we considered the metapopulation biology of *Caenorhabditis* nematodes, which involves  
644colonization of ephemeral habitat patches, rapid population expansion, and then generation of new  
645dispersal morphs (dauers) when the population exhausts its patch (Cutter, 2015; Félix & Braendle,  
6462010; Ferrari et al., 2017; Richaud et al., 2018). If a patch exists for a limited time, patches whose  
647populations expand most rapidly will contribute more descendants to the overall gene pool when the  
648patches expire. Chance genetic differences among patches can strongly influence the rate of  
649population expansion (Figure 10 - figure supplement 1). Within each patch, allele frequencies  
650experience the same forces as in a population of fixed size, but among patches we expect that those  
651founded solely by homozygotes (experiencing no selective deaths from Medeas) will have grown the  
652largest after a few generations, increasing the representation of their alleles in the overall gene pool. As  
653the rarer allele at a locus will be overrepresented in heterozygotes, and most homozygotes will carry  
654only the common allele, this mode of population regulation should increase the frequency of the more  
655common allele. This is a kind of "Haystack Model," well studied in the context of sex ratios, where  
656patchy environments favor a female bias because of the more rapid population expansion it allows  
657(Bulmer & Taylor, 1980; Wilson & Colwell, 1981). These group-selection models are sensitive to  
658parameters, including the number of individuals that found a patch, the number of offspring per parent,  
659and the number of generations of growth within each patch. For each selfing rate, we modeled a  
660situation in which each patch is colonized by four individuals, with sex ratio fixed to exclude its effects  
661but genotypes drawn from the equilibrium frequencies given the allele frequency and selfing rate. The  
662population then grows without any constraints on size until the patch reaches its expiration date, after  
663three generations, at which time we calculate the global allele frequency, summing across the entire  
664collection of patches. Under these conditions, we confirm a strong negative relationship between patch

665heterozygote frequency and patch growth rate. For a single Medea, the effects of growth in a patchy  
666environment on allele frequency are very modest, and the genic selection within each patch allows  
667Medeas to spread despite their costs. With antagonistic Medeas, modeled with equal penetrance, the  
668reduction in heterozygote brood size is greater and the directional effects of single-Medea genic  
669selection are absent. In these conditions, group selection is sufficient to generate a global change in  
670allele frequency (Figure 10C). There are more all-homozygote patches for the more common  
671haplotype, and the result is positive frequency-dependent selection, decreasing the frequency of the  
672less common haplotype. Overall, the joint effects of selfing and group selection on antagonistic Medeas  
673are therefore to eliminate the rarer allele. Selection achieves this result most efficiently at intermediate  
674selfing rates. Although the Medeas stick around at higher selfing rates, they impose very little genetic  
675load, and at very high selfing rates they are entirely neutral.

676

677Figure 10. Selfing reduces the efficacy of Medea elements. (A) Representative allele frequency trajectories of a  
678Medea haplotype under different rates of selfing ( $S$ ) for 100 generations. Population size is 20,000 in each case,  
679the initial Medea allele frequency is 0.05, penetrance is 0.95, and initial genotype frequencies and sex ratios are  
680those expected at neutral equilibrium given the selfing rate. At selfing rates of 95%, elements fail to increase in  
681frequency during 100 simulated generations. (B) Distribution of Medea allele frequencies after 100 generations in  
682populations of size 1000. Each histogram shows the outcome of 250 simulations with initial Medea allele  
683frequency 0.05. Medea alleles are often lost under high selfing rates. (C) Antagonistic haplotypes induce positive  
684frequency dependent selection when selfing rates are intermediate, when populations undergo exponential  
685growth in ephemeral habitat patches, or both. Each boxplot represents the results of 250 simulations of three  
686generations of evolution with competing Medea haplotypes starting from allele frequency 0.2 and 0.8, with initial  
687population size 1000. In the patchy environment, those 1000 individuals are distributed among 250 separate  
688patches, and population growth is unbounded within each. Source code is available from [github](#).

689

690Figure 10 - figure supplement 1. In a single-Medea scenario, genic and group selection affect the frequencies of  
691haplotypes, and Medea-induced deaths generate individual-level selection for suppression.

692

693

694Figure 10 - figure supplement 2. Antagonistic Medeas are subject to drift, genic selection, or positive frequency-  
695dependent selection, depending on the selfing rate. A-C show simulations of antagonistic Medea elements with  
696equal penetrance (0.95), D and E show antagonistic elements with unequal penetrances. (A) At a locus with two  
697haplotypes, each containing Medea elements, allele frequencies change over time as a function of selfing rate  
698( $S$ ). At intermediate selfing rates, positive frequency-dependent selection acts to remove the rarer haplotype from  
699the population. This effect is explained by the disproportionate impact of segregation in heterozygous selfers on  
700the rarer haplotype. Heterozygous selfers kill each haplotype in equal proportions ( $\frac{1}{4}$  of progeny are homozygous  
701for each haplotype), while outcrossers kill each as a function of their population frequency; the result is that the  
702rare allele suffers more from heterozygote selfing. Intermediate levels of outcrossing result in the highest  
703frequency of heterozygous selfer and so the most efficient selection against the rare allele. Under complete  
704selfing ( $S = 1$ ), there are no heterozygotes and Medeas have no effect. Under obligate outcrossing, the  
705haplotypes are subject to drift and show similar dynamics to obligate selfers or outcrossers without Medeas. The  
706pattern in obligate outcrossers reflects the maternal-zygotic interaction character of the Medeas. While the  
707Medeas kill homozygotes, generating overdominance and heterozygote excess, this occurs only in heterozygous  
708mothers. The fitness cost to these mothers represents underdominance, favoring homozygotes, and the over- and  
709under-dominance balance each other. This figure shows the results of representative simulations with fixed  
710population size of 20,000 breeding individuals, initial frequencies 0.2 and 0.8, and 95% penetrance of each of the

711two Medeas. **(B)** Heterozygote frequencies (from the simulations shown in A) vary over the simulations. Obligate  
712outcrossers maintain substantially elevated heterozygote frequencies. **(C)** Surprisingly, despite the strong  
713selection against homozygotes and the elevated frequency of heterozygotes in obligate outcrossers, allele  
714frequencies evolve with drift-like dynamics. Results of 200 simulations are shown, each with a fixed population  
715size of 1,000 breeding individuals and initial minor allele frequency of 0.2. The left panel illustrates the spread of  
716allele frequencies across simulations in the case of antagonistic Medeas, and the right shows results with no  
717Medeas. In both cases, frequencies drift and some populations lose the lower-frequency allele. After 200  
718generations, the allele frequency variances are not different between the two cases (Levene's test,  $p = 0.13$ ). **(D)**  
719When a strong Medea element (penetrance 0.95) invades a population carrying a weaker Medea (penetrance  
7200.6), its ability to displace the weak Medea depends on the selfing rate ( $S$ ). At low selfing rates, genic selection  
721favoring the strong Medea drives it to high frequency. At intermediate and high selfing rates, positive frequency-  
722dependent selection favors the weaker resident Medea, and the strong haplotype is removed from the population.  
723The figure shows the results of representative simulations with fixed population size of 20,000 breeding  
724individuals, initial frequencies 0.2 and 0.8, and penetrances chosen to represent those of the *C. tropicalis*  
725JU1373 haplotype (0.95) and NIC58 haplotype (0.6). **(E)** Heterozygote frequencies (from the simulations in D)  
726evolve, with heterozygosities transiently elevated above 0.5 during sweeps. Source code is available from [github](#).

## 727Discussion

728We have brought genetic and genomic resources to the most recently discovered androdioecious  
729*Caenorhabditis* species, *C. tropicalis*, facilitating studies of the varied effects of mating system  
730transitions on genomes and metapopulation genetics, and their interactions with ecology, as well as  
731comparative quantitative genetics. Aided by a chromosome-scale reference genome, our data confirm  
732that *C. tropicalis* is indeed the most genetically homogeneous of the three selfers, on average, which  
733implies a correspondingly high rate of selfing, low rate of effective recombination, and small effective  
734population size. We also show that the average obscures extreme variance in the distribution of  
735genomic diversity. This finding mirrors recent findings for *C. elegans* and *C. briggsae* (Lee et al., 2020).  
736Within two highly divergent regions, we find some striking biology: three Medea gene-drive systems  
737segregating in a single cross. Genetically similar Medea elements have also been found in *C. elegans*  
738(Ben-David et al., 2017; Seidel et al., 2008), but we and others have now shown that they are  
739especially common in *C. tropicalis* (Ben-David et al., 2020). They thus represent a potent and  
740surprisingly widespread form of genetic incompatibility underlying outbreeding depression, and a  
741potential cause of the species' high effective selfing rate. Comparative analysis will be furthered by  
742more extensive sampling, in time, and in geographic and genomic space, for all selfing species. This  
743project will also benefit from inclusion of the closely related sister-species of *C. tropicalis*, *C. wallacei*  
744(Félix et al., 2014), as in the case of *C. briggsae* and its outcrossing sister *C. nigoni* (Yin et al., 2018).

## 745Selfing and population genetics

746Our early view of selfing *Caenorhabditis* species was of widespread, weedy lineages depauperate of  
747genetic diversity relative to their outcrossing ancestors. This view, limited by mostly opportunistic  
748sampling of isolates, and sequencing technology, was based on the apparent global expansion of *C.*  
749*elegans* associated with human activity (Andersen et al., 2012). Better sampling has led to a more  
750complete picture of strong population and genomic structure for all three species (Crombie et al., 2019;  
751Thomas et al., 2015), though tropical areas remain particularly undersampled.

752

753Most recently, a large survey of *C. elegans* genomes, including 15 assembled from long-reads, found  
754regions of high-diversity spanning up to 20% of the reference genome (Lee et al., 2020). Similar  
755heterogeneity was found in a smaller sample of *C. briggsae* genomes. This finding was presaged by  
756efforts to build a complete genome for the divergent Hawaiian isolate CB4856 using second generation

757sequencing (Thompson et al., 2015), but has been greatly enabled by more contiguous assemblies that  
758circumvent the reference mapping bias plaguing study of all genetically diverse species. A promising  
759hypothesis for the presence of hyperdivergent regions in genomes is the action of balancing selection  
760across the species' range, leading to preservation of some of the presumably abundant genetic  
761diversity of outcrossing ancestral species. This hypothesis is supported in *C. elegans* by the enrichment  
762of genes encoding environmentally responsive sensory factors, which are themselves enriched for  
763differential expression and quantitative trait loci for response to microbes isolated from natural habitats.  
764Alternative hypotheses include introgression, a common source of islands of divergence in other animal  
765taxa (Hedrick, 2013), but the presence of multiple distinct divergent haplotypes in *C. elegans* (Lee et  
766al., 2020) argues strongly against it. Balancing selection can act at many levels, from local adaptation  
767via directional selection, at the global scale, to various frequency-dependent phenomena at the local  
768scale, and overdominance at the molecular scale. While some environmental associations clearly play  
769a role in the biogeography of all three androdioecious species (Crombie et al., 2019; Cutter et al., 2006;  
770Dolgin et al., 2008; Félix & Duveau, 2012; Ferrari et al., 2017; Kiontke et al., 2011; Prasad et al., 2011;  
771Thomas et al., 2015), they seem not to be a definitive factor in structuring divergent haplotypes. The  
772high migration rate that comes with the microscopic nematode form, coupled with the reproductive  
773assurance afforded by selfing, may effectively counter adaptation to global environmental variation.

774  
775Taxa with mixed selfing and outcrossing, very large population sizes, and broad, diverse ranges, may  
776occupy a population genetic space particularly well-suited to the detection and localisation of balancing  
777and local selection. The ability to detect targets of strong balancing selection scales with total  
778population size (assuming symmetric migration) and recombination rate, and the homogenising effect  
779of partial selfing is to increase the signal of balanced peaks against a background continually swept of  
780diversity by indirect selection (Charlesworth et al., 1997; Nordborg et al., 1996). The particularly high  
781selfing rate of *C. tropicalis* may be especially favorable, and its preferred habitat may also allow for  
782more rapid evolution than the temperate-dwelling *C. elegans*. Detection of the targets of local  
783adaptation depends on population size, migration rate and selection intensity (Charlesworth et al.,  
7841997), which are all unknowns. Although Medea factors may have the potential to generate balancing  
785selection, it is clear that most divergent regions between NIC58 and JU1373 do not carry Medea  
786factors, or, at least, not ones that are active independent of environment and genetic background.  
787These, and other highly divergent haplotypes in *C. tropicalis*, may harbor loci under positive selection in  
788different conditions, a conjecture made more plausible by the analysis of divergent gene content in *C.*  
789*elegans*.

## 790Population dynamics of Medea elements

791Medea elements like those we discovered in *C. tropicalis* will reliably spread in randomly mating  
792populations (Wade & Beeman, 1994). Highly penetrant killing of homozygous larvae is an exceptionally  
793potent selective mechanism to drive allele frequency change, and in our short RIL construction  
794pedigree we saw two haplotypes sweep nearly to fixation. In nature, things are likely quite different, as  
795both the mating system and natural history of *C. tropicalis* conspire to render Medea loci selectively  
796inert. *C. tropicalis* shows strong geographic structure, presumably exacerbated by habitat  
797fragmentation, such that encounters between Medea and sensitive haplotypes may be rare. Upon an  
798encounter between a hermaphrodite and a male, outcrossing rates, measured under benign laboratory  
799conditions, are relatively low on average. When divergent isolates do cross, Medeas may find  
800themselves inactive due to dependence on genetic background, including mitochondrial genotype, and,  
801potentially, environmental factors. Most importantly, *C. tropicalis* reproduces primarily by self-  
802fertilization, and gene-drive elements are unable to gain traction when heterozygotes are infrequent. At

803the same time, the patchy, ephemeral microhabitat of *C. tropicalis* – rotting fruits and flowers on the  
804forest floor – provides a perfect substrate for group selection. Small numbers of dispersing larvae  
805colonize each patch and undergo exponential population growth for a small number of generations.  
806Although Medeas will increase in frequency in patches with heterozygotes, population growth in  
807patches without heterozygotes can be so much greater as to overwhelm the countervailing effects of  
808Medeas on allele frequency.

809

810The patterns we observe on chromosome V implicate tightly linked Medeas, one on each of the  
811alternative haplotypes, a phenomenon also discovered by Ben-David et al. (2020). Surprisingly, we  
812found by simulation that antagonistic Medeas do not generate balancing selection, at least under the  
813scenarios modeled. They nevertheless impose a strong segregation load under outcrossing, which  
814should select for suppressors. The antagonistic Medeas themselves evolve by drift at high selfing rates,  
815and at intermediate selfing rates frequency-dependent selection eliminates the rarer haplotype. At the  
816same time, the chromosome V Medeas occur on ancient haplotypes, evidenced by extreme divergence  
817between NIC58 and JU1373 (Figure 5; Figure 8). These haplotypes may encode unique toxin-antidote  
818pairs that arose independently, or they may encode toxin-antidote pairs that co-evolved from a common  
819ancestor but are no longer cross-compatible. Competition among driver haplotypes is known to occur  
820for Segregation Distorter in *D. melanogaster*, but in that case, driver haplotypes compete for slots in a  
821balanced equilibrium with non-drivers (C. L. Brand et al., 2015; Presgraves et al., 2009). Our  
822simulations raise questions about whether antagonistic Medeas play a role in the ancient balancing  
823selection at the chromosome V locus, and we note that the two haplotypes are sufficiently different in  
824gene content that effects on other phenotypes are likely.

## 825Molecular mechanisms of Medea-mediated gene drive

826The Medea factors we have discovered in *C. tropicalis* are analogous to the *sup-35/pha-1* maternal-  
827effect driver (Ben-David et al., 2017) and the *zeel-1/peel-1* paternal-effect driver in *C. elegans* (Seidel et  
828al., 2008, 2011), and four maternal-effect Medea drivers in *Tribolium* (Beeman et al., 1992; Beeman &  
829Friesen, 1999). Additionally, several other *C. tropicalis* Medea factors have been independently  
830identified and characterised by Ben-David et al. (2020). Similar inheritance patterns have also been  
831reported for two loci in mice (Peters & Barker, 1993; Weichenhan et al., 1996, 1998; Winking et al.,  
8321991). The causal genes underlying the JU1373 and NIC58 Medeas remain to be identified, but likely  
833include one or more of the multiple genes unique to these haplotypes, as seen for *C. elegans* where  
834toxin and antidote functions are encoded by genes present on the killing haplotype and absent (or  
835pseudogenized) on the non-killing haplotype (Ben-David et al., 2017; Seidel et al., 2008). Similarly, the  
836single *Tribolium* Medea element whose genetic basis is known maps to a transposable element  
837insertion absent from non-Medea haplotypes (Lorenzen et al., 2008). A common pattern emerging from  
838these systems is that maternal-effect drivers (and the single example of a paternal-effect driver) are  
839encoded by dispensable genes with dedicated functions, rather than genes acquiring toxin or antidote  
840activity while retaining an ancestral non-drive function.

841

842Why are Medeas so prevalent in *C. tropicalis* (and to some extent in *C. elegans* and *C. briggsae*) but  
843mostly absent elsewhere? One option is ascertainment bias – maybe similar elements are  
844taxonomically more widespread, but we simply haven't looked for them.

845

846A second option is that mechanisms of translational control in the *Caenorhabditis* germline may make it  
847easy for Medeas to arise. Early embryogenesis in *Caenorhabditis* is largely controlled by maternal  
848regulators, and a common expression pattern for these regulators is ubiquitous expression of mRNA in

849the oocyte but no translation until embryogenesis (Evans & Hunter, 2005; Robertson & Lin, 2015). A  
850gene whose protein is generally cytotoxic could become a maternal-effect toxin by acquiring the  
851(common) regulatory elements specifying this expression module. If such a mutation occurs in tight  
852linkage to an incidental zygotically expressed antidote, it creates a Medea. Variation among species in  
853maternal gene regulation may therefore be relevant to variation in the mutational flux of novel Medeas.  
854

855The antidote-first scenario, in which Medeas arise from sequential mutations on a single haplotype,  
856may be facilitated by high rates of selfing. In populations with high selfing rates, alternate haplotypes  
857rarely encounter one another in heterozygotes. Each can then evolve as though in allopatry with the  
858other. If balancing selection preserves the alternate haplotypes for a long time, they may incidentally  
859accumulate antidotes and toxins that are neutral in their own backgrounds (Seidel et al., 2008).

860

861Toxin-first evolution is also possible, aided by metapopulation structure and selfing. Partially penetrant  
862toxins might arise in a background lacking an antidote but become locally fixed, despite their  
863deleteriousness, due to the tiny effective population size and inefficient purifying selection of a local  
864selfing population. If outcrossing is rare, toxin-free haplotypes will not be re-introduced or decoupled by  
865recombination and instead, the population might restore its fitness via compensatory evolution of an  
866antidote; if the antidote is linked to the toxin, a Medea is born.

867

868Whether the toxin-antidote elements in *C. tropicalis* and *C. elegans* arose before the transition to selfing  
869is unclear, although the level of divergence between opposite haplotypes at Medea loci is suggestive of  
870sampling from outcrossing ancestors. A closer examination of gonochoristic species in the Elegans  
871group is needed to determine whether toxin-antidote elements are specific to, or quantitatively different  
872in, selfers. It will also be interesting to see the strong outbreeding depression in other taxa with mixed-  
873mating, such as the “cryptic biological species” complexes in arctic *Draba* (Grundt et al., 2006),  
874dissected genetically.

### 875Strategies to combat Medea factors

876Gene drive systems create a selective environment favoring the evolution of suppressors. Suppressors  
877of meiotic and gametic drive have been well documented in many species, especially when obligate  
878outcrossing continually exposes individuals to the costs of drive (Courret et al., 2019; Lindholm et al.,  
8792016; Lyttle, 1991; T. A. R. Price et al., 2019). Suppressors of drive in selfing species are more rare,  
880which has been interpreted as evidence that many drivers in selfing species did not evolve as drivers  
881per se but instead evolved through non-drive mechanisms (Sweigart et al., 2019), such as balancing  
882selection maintaining alternate homozygous genotypes (Seidel et al., 2008). Our data show that  
883unlinked modifiers affect Medea activity in *C. tropicalis*, though whether these modifiers evolved as  
884suppressors is equivocal. In the case of maternal-effect drivers, mitochondrial suppressors are special:  
885selection for mitochondrial suppressors may be especially strong because mitochondria cannot  
886segregate away from drivers via inheritance in sperm. This selective environment may explain why two  
887of the Medeas we discovered in *C. tropicalis* (the chromosome III JU1373 Medea and the chromosome  
888V NIC58 Medea) were differentially active according to mitochondrial genotype – the mitochondrial  
889genotypes non-permissive for Medea drive may have evolved as suppressors. Alternatively, Medeas  
890may have arisen in permissive mitochondrial backgrounds, with little selection for or against alternate  
891mitochondrial genotypes. Ultimately, our data provide little conclusive evidence that the Medea loci  
892experience selection in nature that is due to their drive activity.

893

894The data suggest that crosses between geographically distant *C. tropicalis* isolates will typically reveal  
895multiple Medea loci (Figure 9). Segregation of multiple Medeas magnifies the cost of outcrossing and  
896reduces the possibility of suppression by a common molecular mechanism, in a manner analogous to  
897the role of multidrug therapy in preventing the evolution of drug resistance. The difficulty that organisms  
898face in evolving suppressors to multiple drive elements at once has emerged as an important  
899consideration for gene drive strategies for controlling disease vectors (Burt, 2003; Champer et al.,  
9002018). In such cases, organisms can adapt by altering their population biology, increasing their rates of  
901inbreeding and selfing (Bull, 2017; Bull et al., 2019; Drury et al., 2017), and thus reducing the  
902heterozygosity required for all gene drive activity.

903

904The costs of selfing as a defense against gene drive are inbreeding depression; reduced ability to adapt  
905to new conditions; and reduced genetic variation and hence niche breadth. Androdioecious  
906*Caenorhabditis* appear to have mechanisms for dealing with each of these costs. Selfing  
907*Caenorhabditis* are typically found in nature as totally inbred lines, consistent with having purged  
908recessive deleterious variants in their history (Anderson et al., 2010; Richaud et al., 2018). While  
909outcrossing plays an important adaptive role in selfing *Caenorhabditis* (Chelo et al., 2019; Morran et al.,  
9102009; Teotónio et al., 2006; Teotonio et al., 2012), populations can transiently increase male frequency  
911to achieve adaptation before returning to a primarily selfing mode of reproduction (Anderson et al.,  
9122010; Shi et al., 2017). Finally, the preservation of genetic diversity at large numbers of ancient  
913haplotypes by balancing selection allows these species to occupy a wide range of habitats despite low  
914levels of baseline genetic variation (Lee et al., 2020).

915

916We have shown that *C. tropicalis* harbors abundant heritable variation in outcrossing rate, with  
917nondisjunction, male mating propensity, and hermaphrodite mating propensity all providing avenues for  
918genetic fine-tuning of the outcrossing rate. Other data also show that *C. tropicalis* has mechanisms that  
919promote selfing over outcrossing. Ting et al. (2014) found that *C. tropicalis* hermaphrodites are uniquely  
920resistant to the deleterious effects of interspecific matings, and they interpret their findings as evidence  
921for reduced activity of sperm guidance cues in *C. tropicalis* hermaphrodites. Shi et al. (2017) showed  
922that male longevity is reduced in *C. tropicalis* when male pheromone is present, creating a negative  
923feedback that tamps down male frequencies, but not in obligately outcrossing *Caenorhabditis* species.  
924These findings are consistent with selection favoring high selfing rates in *C. tropicalis*.

925

926Selfing is often considered a factor that favors the evolution of incompatibilities and outbreeding  
927depression, just as the independent evolution of species or subspecies often leads to incompatibilities  
928revealed by hybridization (Fishman & Sweigart, 2018; Maheshwari & Barbash, 2011; Presgraves,  
9292010). Selfing reduces the effective recombination rate, allowing unlinked loci to evolve together. When  
930outcrossing reshuffles these co-evolved loci, it creates new combinations of alleles untested by  
931selection. Incompatibilities between these alleles manifest as outbreeding depression (equivalently,  
932rearrangements can fix within selfing lineages, rendering outbred progeny deficient). Our findings  
933suggest causation running in the opposite direction should also be considered. Incompatibilities in *C.*  
934*tropicalis* appear to mostly represent interactions between tightly linked loci acting in different  
935individuals (mothers and offspring), rather than interactions between unlinked loci; thus, the resulting  
936outbreeding depression is mostly not mediated by recombination. *C. tropicalis* can escape these  
937incompatibilities and restore fitness by inbreeding. Thus, in contrast to the usual pattern of selfing  
938leading to incompatibility, in this species incompatibility may also lead to increased selfing.

939



## 941Methods

### 942Strain maintenance

943Strains were maintained using standard protocols for *C. elegans* (Brenner, 1974; Stiernagle, 2006),  
944with the addition of 1.25% agarose to NGM-agar (NGMA) plates to discourage burrowing, and a 25C  
945incubation temperature. This temperature is characteristic of substrate temperatures where we have  
946collected *C. tropicalis*, and is the standard rearing temperature in previous work on this species  
947(Gimond et al., 2013).

### 948Genome sequencing

949Long-read data for NIC58 and JU1373 were around 250x expected coverage, given an estimated  
950genome size of roughly 80 Mb (Fierst et al., 2015), from a PacBio Sequel at the Duke University Center  
951for Genomic and Computational Biology. DNA was extracted from twelve 10 cm NGMA plates of  
952nematodes spotted with OP50 using the Qiagen MagAttract HMW DNA kit as per Lee et al. (2020).  
953

954JU1373 and NIC58 short-read data were around 25x and 40x expected coverage 100 bp paired-end  
955reads (TruSeq libraries, HiSeq 2000, NYU Center for Genomics and Systems Biology Genomics Core),  
956and another 40x coverage for NIC58 (150 bp paired-end reads, TruSeq library, NovaSeq6000,  
957Novogene).

958

959We sequenced 129 RILs from the cross between JU1373 and NIC58 to a median depth of 2.1x  
960(NextEra libraries using the protocol of Baym et al. (2015), NextSeq 500, paired-end 75 and 150 bp  
961reads, NYU Center for Genomics and Systems Biology Genomics Core). DNA was isolated by  
962proteinase-K digestion followed by phenol/chloroform/isoamyl alcohol purification.  
963

964An additional 22 wild isolates were sequenced to a median depth of 29x (NextEra libraries, NextSeq  
965500, paired-end 75 bp reads; or BioO libraries, HiSeq 2000, paired-end 100 bp reads; NYU Center for  
966Genomics and Systems Biology Genomics Core). DNA was extracted by salting out (Sunnucks &  
967Hales, 1996). Isolates and associated metadata are in Figure 3 - source data 1.

968

969All sequencing reads used in this project are available from the NCBI Sequence Read Archive under  
970accession [PRJNA662844](https://www.ncbi.nlm.nih.gov/PRJNA662844).

### 971Genome assembly

972Our reference genome is NIC58. We generated initial assemblies for evaluation with genetic linkage  
973data, including a Canu hybrid assembly (Koren et al., 2017) and long-read only assemblies from flye  
974(Kolmogorov et al., 2019), ra (Vaser & Šikić, 2019) and wtgbg2 (Ruan & Li, 2019). All were initially run  
975with default parameters. Flye produced a highly contiguous assembly with this data, and initial genetic  
976evaluation showed few errors (interchromosomal chimeras were detected for all other assemblers), so  
977we varied parameters (minimum overlap length 4-10 kb, initial assembly depth 40-180x) and selected  
978the two most contiguous assemblies for closer evaluation (the genetically concordant assembly used -  
979m 10 kb --asm-coverage 120x). A draft assembly for JU1373 was made with flye using default  
980parameters (44 contigs and scaffolds, NG50 4.2 Mb, 81 Mb span). Both assemblies were polished with  
981short-reads using Pilon (-fix bases mode) before further use (Walker et al., 2014).  
982

983 Mitochondrial genomes were initially assembled from long reads mapping to contigs identified as  
984 partially homologous to *C. elegans* sequence. *De novo* assemblies using Unicycler (Wick et al., 2017)  
985 to produce a polished circular sequence showed homology to all 12 *C. elegans* proteins for both NIC58  
986 and JU1373, but total length and sequenced identity were sensitive to input read length (using all data,  
987 or only reads of length 10–15 kb, which spans the range of full-length *Caenorhabditis* mitochondrial  
988 genomes in GenBank) and mapping quality. We instead used fragmented, high-coverage contigs from  
989 Illumina *de novo* assemblies (Platanus 1.2.4; (Kajitani et al., 2014)) with homology to the long-read  
990 assemblies as bait to extract short reads for reassembly, which produced single sequences of length  
991 13565 and 13091 bp for NIC58 and JU1373, respectively. After circular polishing with long-reads  
992 (Unicycler), sequences were 14394 and 14027 bp. We rotated these with five copies of the *C. elegans*  
993 mitochondrial genomes to optimise linear homology using MARS (Ayad & Pissis, 2017).

#### 994 **Genetic map construction**

995 RILs were derived by crossing a NIC58 male and JU1373 hermaphrodite, and inbreeding the F<sub>2</sub>  
996 offspring of a single F<sub>1</sub> hermaphrodite for 10 generations by selfing. We used the RIL data to evaluate  
997 assemblies based on interchromosomal consistency and concordance between genetic and physical  
998 marker order. A SnakeMake pipeline (Köster & Rahmann, 2012) implementing this procedure is on  
999 [github](#).

1000

1001 Using short-read mapping to the NIC58 assemblies, we called variants distinguishing the parental lines,  
1002 filtered them to homozygous diallelic SNVs (depth within 1/3 of the median, > 10 bp from an  
1003 insertion/deletion, quality > 50, then removing any SNVs in 20 bp windows with more than one SNV),  
1004 and genotyped RILs at the remaining sites (H. Li, 2011; H. Li et al., 2009; H. Li & Durbin, 2009;  
1005 Vasimuddin, Md et al., 2019).

1006

1007 Parental ancestry was inferred by HMM (Andolfatto et al., 2011), sampling one variant per read, with  
1008 transition probabilities defined by homozygous priors, recombination rate ( $r$  = per base pair rate given  
1009 an expected 6 recombination events per RIL genome), physical distance between markers in the  
1010 reference genome ( $d$ ) and a scaling factor ( $r_{fac} = 10^{-11}$ ), parameterised as  $10^{-r*d*r_{fac}}$ , and emission  
1011 probabilities set by parental genotyping error rate ( $10^{-4}$ ) and read base quality scores. Markers for map  
1012 construction were defined by filtering on posterior probability > 0.5, binning up to 50 SNVs, and merging  
1013 the sparse RIL marker inferences, interpolating missing positions across consistent uniparental flanking  
1014 bins. Bins with both parental genotypes were considered as missing data.

1015

1016 Marker filtering and map construction was carried out in R/qtl (Broman et al., 2003). After dropping  
1017 identically informative genotypes, two lines that were outliers for heterozygosity, and one of each of  
1018 eight pairs of lines with >99% similarity, linkage groups (LGs) were formed (maximizing the number of  
1019 markers in six LGs), and markers were ordered within LGs by likelihood from 100 iterations of greedy  
1020 marker ordering. Where genetic and physical ordering conflicted, the physical order was tested by  
1021 likelihood and accepted if the change in LOD was > -1. Taking the genetic data as ground truth, we  
1022 compared assemblies on the number of sequences spanning more than one LG, and on the number  
1023 and sum of negative LOD scores for any remaining discordance in within-LG genetic/physical marker  
1024 order.

1025

## 1026Genome orientation and scaffolding

1027On the above metrics, we selected a flye assembly, spanning 81.83 Mb in 36 contigs and scaffolds >  
102820 kb with an N50 of 4.795 Mb (half the expected genome size of 80 Mb was in sequences of at least  
1029this length). The genetic map based on this assembly incorporated 33 of these sequences and  
1030spanned 81.3 Mb. We then did two rounds of manual stitching, considering only junctions with  
1031estimated genetic gaps of 0 cM. First, we accepted 10 joins where sequences from another assembly  
1032spanned a junction (>5 kb of MQ=60 alignment on either flank; minimap2 (H. Li, 2018)). Second, we  
1033accepted 8 joins where at least one read consistent with the genetic orientation spanned a junction  
1034(alignment >2 kb of MQ>20 on each 50 kb flank; minimap2). We took the consensus sequence  
1035(bcftools), or in two cases the read sequence, and converted the now fully-oriented assembly of 15  
1036sequences into pseudochromosomes, with 50 bp N gaps at the remaining junctions. Chromosomes  
1037were named and oriented based on *C. elegans* homology, by summing aligned lengths per  
1038chromosome and strand (minimap2 -x asm20 mapping quality >30). Chromosome preference was  
1039unequivocal (>60-fold bias toward a single homolog in all cases). Strand preference was relatively  
1040strong for chromosomes II and X (>3.2-fold bias), but less so for the others (1.8-fold bias for IV, 1.4 for  
1041I, 1.3 for III, and 1.1 for V), from 0.400-1.9 Mb of aligned sequence per homologous chromosome. The  
1042inferred orientations were consistent with strand bias from 1:1 orthologs in chromosome centers in all  
1043cases except chromosome I. Finally, we did one further round of short-read polishing (pilon -fix bases  
1044mode, making 5044 changes), and re-estimated the genetic map.

## 1045Annotation

### 1046Mixed-stage RNA preparation

1047We collected three samples each for NIC58 and JU1373: well-fed mixed-staged (L1-adults), well-fed  
1048male-enriched, and starved (including dauers) plates. Strains were passaged by chunking every two  
1049days to maintain a well-fed mixed-stage population. Some plates were allowed to starve, and the  
1050presence of dauer larvae along with other developmentally arrested larvae was confirmed by visual  
1051inspection. Crosses were set up on single-drop OP50-seeded plates with 15-20 males and a few  
1052hermaphrodites to establish a male-enriched population. Following successful mating, worms were  
1053chunked to 10 cm OP50-seeded plates for sample collection.

1054

1055Each sample was collected from one 10 cm plate, flash frozen in 100 µl S-Basal in liquid nitrogen, and  
1056stored at -80C until extraction. Total RNA was extracted using Trizol reagent (Invitrogen) following the  
1057manufacturer's protocol, except that 100 µl acid-washed sand (Sigma) was added during the initial  
1058homogenization step. RNA was eluted in nuclease-free water, purity was assessed by Nanodrop  
1059(ThermoFisher), concentration was determined by Qubit (ThermoFisher), and integrity was assessed  
1060by Bioanalyzer (Agilent). Following quality control, 1.5 µg of total RNA from each sample was pooled,  
1061further purified using the RNA MinElute Cleanup kit (Qiagen), and again subject to the above quality  
1062control analyses.

1063

### 1064Library preparation and sequencing

1065RNAseq libraries for JU1373 and NIC58 were prepared simultaneously from mRNA isolated from 1 µg  
1066of pooled total RNA using the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England  
1067Biolabs). RNA fragmentation, first and second strand cDNA synthesis, and end-repair processing were  
1068performed with the NEBNext Ultra II RNA Library Prep with Sample Purification Beads (New England  
1069Biolabs). Adapters and unique dual indexes in the NEBNext Multiplex Oligos for Illumina (New England  
1070Biolabs) were ligated, and the concentration of each library was determined using Qubit dsDNA BR

1071Assay Kit (Invitrogen). Libraries were pooled and qualified by Bioanalyzer 2100 (Agilent; Novogene,  
1072CA, USA), and 150 bp paired-end reads were sequenced on a single Illumina NovaSeq 6000 lane.

1073

#### 1074Gene prediction

1075We identified repetitive sequences in the NIC58 and JU1373 genomes de novo using RepeatModeler  
1076(Smit et al., n.d.) and classified these using the RepeatClassifier tool from RepeatModeler and the  
1077Dfam database (Hubley et al., 2016). We removed unclassified repeats and soft-masked the genome  
1078assemblies with RepeatMasker using the classified repeat library. We aligned short RNAseq reads to  
1079the soft-masked genomes with STAR in two-pass mode (Dobin et al., 2013), and used the BRAKER  
1080pipeline to annotate genes (Hoff et al., 2019). We extracted protein sequences from the BRAKER  
1081annotation using the getAnnoFasta.pl script from AUGUSTUS (Stanke et al., 2006), and assessed  
1082biological completeness using BUSCO (Seppey et al., 2019). We annotated mitochondrial genomes by  
1083homology using the MITOS2 server (Bernt et al., 2013) and Prokka (Seemann, 2014).

### 1084Population genetics

#### 1085Variant calling

1086A SnakeMake pipeline implementing variant calling and filtering is available from [github](#) (Köster &  
1087Rahmann, 2012). In brief, we mapped reads to the NIC58 reference genome with bwa mem2  
1088(Vasimuddin, Md et al., 2019), aligned and normalized indels with bcftools (H. Li, 2011), called variants  
1089jointly with GATK (DePristo et al., 2011; McKenna et al., 2010), and hard filtered diallelic SNVs (median  
1090absolute deviation in total depth < 99th percentile, QD > 4, MQ > 30, BaseQRankSum > -3,  
1091ReadPosRankSum > -4, SOR < 5). We also applied per-sample depth filtering (local depth in 1 kb  
1092windows < 2x against a LOESS polynomial fit for each chromosome, span=0.33), keeping SNVs in  
1093windows where at least 22/24 samples passed. A total of 880,599 diallelic SNVs were called, 794,676  
1094passed filtering (genotype set 1), and we used the fully homozygous subset of these with no missing  
1095data, comprising 397,515 SNVs (genotype set 2).

### 1096Population structure

1097Principal component analysis was carried out on the additive genetic relationship matrix (base R  
1098`prcomp`) constructed from homozygous diallelic SNVs with no missing data (genotype set 2).

### 1099Divergent regions

1100We thresholded divergent regions using kernel density smoothing (Duong, 2020) of the empirical  
1101distribution of  $\theta_w$  across genomic windows (10 kb), taking the first positive value of the first derivative,  
1102after the minimum, as the threshold value. Regions were enumerated based simply on contiguous runs  
1103of the sign of the second derivative of  $\theta_w$ , that is, all local peaks in nucleotide diversity are treated as  
1104independent. This makes the unrealistic assumptions of uniform ancestral diversity and effective  
1105recombination, and is sensitive to sample size and binning. Deeper and broader population genetic  
1106data will be required to obtain more confident estimates of the number, size and local structure of  
1107divergent regions, ideally with multiple high-quality genome assemblies to minimize confounding by  
1108reference mapping bias.

### 1109Statistical analysis, data wrangling, plotting

1110We used R (R Core Team, 2018) with packages data.table (Dowle & Srinivasan, 2019), dglm (Dunn &  
1111Smyth, 2016), dplyr (Wickham et al., 2020), ggmap (Kahle & Wickham, 2013), ggplot2 (Wickham,

11122016), ggh4x (T. van den Brand, 2020), ggrepel (Slowikowski, 2020), lme4 (Bates et al., 2015), and  
1113tidyR (Wickham & Henry, 2020).

#### 1114Mating trials among isolates and RILs

1115Mating trials were initiated with one L4 hermaphrodite and one L4 male worm on a 6 cm NGM agarose  
1116plate seeded with 50 µL of OP-50 *E. coli*. Plates were scored 72 hours later, with success defined as  
1117the presence of multiple males in the F<sub>1</sub> generation.

1118

1119We scored hermaphrodite cross probability in RILs by crossing NIC58 males to L4 RIL hermaphrodites.  
1120RIL trials ranged from 16 to 75 in number, with a median of 30, and took place over 116 days. A total of  
1121338 JU1373 and 412 NIC58 control crosses were done on 107 of these days.

1122

1123To estimate equilibrium male frequency, we scored sex ratio after 10 generations of passaging at large  
1124population size. Three L4 hermaphrodites and 5 L4 males were placed on a 6 cm agarose plate. Three  
1125days later, 3 mL of M9 buffer was pipetted onto the plate and 50 µL of worms was transferred to a 10  
1126cm plate. 50 µL of worms was subsequently transferred to a 10 cm agarose plate every ~72 hours for  
112710 generations, at which point a sample of 267 ± 27 worms were sexed per strain. We performed three  
1128replicates of this passaging experiment.

1129

1130Phenotypes for RIL QTL mapping were best linear unbiased predictions (BLUPs) from a binomial linear  
1131mixed-effects model (R package lme4; (Bates et al., 2015)). Mapping in R/qtl (Broman et al., 2003)  
1132used a ‘normal’ model, and 1000 permutations of the phenotype values to establish genome-wide  
1133significance. The variance explained by the single significant QTL was estimated from variance  
1134components by refitting the linear model to the raw data with a random effect of genotype within RIL.

#### 1135Genetic analysis of Medea activity

1136We used a standardized assay of the proportion of F<sub>2</sub> embryos that develop to adulthood according to  
1137wild-type schedule. P<sub>0</sub> males and hermaphrodites were paired as L4s, and the following day each  
1138hermaphrodite that bore a copulatory plug was transferred to a fresh plate to lay embryos. Two days  
1139later, when these embryos had developed to L4 stage, we isolated F<sub>1</sub> hermaphrodites overnight. The  
1140following day, hermaphrodites were singled to new plates and left to lay embryos for 8 hours. The  
1141hermaphrodites were then removed and the embryos on each plate counted. Three days later, when  
1142wild-type animals have reliably reached adulthood, we counted the adults on each plate by picking. In  
1143some cases, slow-developing animals that had reached L3 or L4 were counted separately. The majority  
1144of Medea-affected animals arrest as L1s and are difficult to see, so we typically estimated the number  
1145of arrested larvae as the number of embryos initially observed minus the number of adults counted  
1146three days later. In a small number of broods (~3%), the count of progeny at adulthood exceeded the  
1147count of embryos laid (by at most two extra adults, from broods containing ~35-55 total embryos laid).  
1148We made the assumption that this discrepancy reflected undercounting of embryos rather than  
1149overcounting of adults, given that embryos are hard to see. Thus, for such broods, we adjusted the  
1150embryo count upward to match the count of adults. All conclusions are robust to this adjustment.

1151

1152Because of the low mating efficiency of many *C. tropicalis* genotypes, matings did not always produce  
1153cross progeny. To distinguish self and cross progeny, we employed several approaches. In some  
1154experiments, we depleted hermaphrodites of sperm by transferring them to fresh plates on each of the  
1155first four days of adulthood, until they ceased reproduction. These sperm-depleted hermaphrodites

1156could then be crossed to males, and resulting progeny inferred reliably to be cross offspring. This  
1157method is not suitable for all experiments because the sperm-depleted hermaphrodites have small  
1158broods and generally show age-associated decrepitude. As an alternative, we developed visible marker  
1159strains that allow us to distinguish self and cross progeny. We isolated a spontaneous Dumpy mutant in  
1160the JU1373 background and established strain QG2413, *Ctr-dpy* (*qg2*). Control experiments confirmed  
1161that this semidominant mutation allows for clean discrimination between Dpy and semi- or non-Dpy  
1162animals, and that the mutation is unlinked to the Medea loci on chromosomes III and V. Next, we  
1163generated a NIC58 derivative carrying a fluorescent reporter. Strain QG3501 (*qgl/s5*) carries pCFJ104  
1164[*Pmyo-3::mCherry::unc-54utr*] (Frøkjaer-Jensen et al., 2008). The transgene was introduced by  
1165microinjection into NIC58, integrated by UV, and backcrossed to NIC58 seven times. These animals  
1166express bright red fluorescence in muscle, visible under the dissecting scope from mid embryogenesis.  
1167Control experiments show that this transgene is unlinked to the Medea loci. To test for maternal-effect  
1168killing by the NIC58 haplotype on chromosome V, we generated strain QG4249, which carries the *qgl/s5*  
1169transgene and the NIC58 mitochondrial genome and is homozygous NN at the chromosome III locus  
1170and JJ at the chromosome V locus. This strain was generated by crossing QG2514 males and QG3501  
1171hermaphrodites and recovering a rare F<sub>2</sub> adult that was homozygous JJ at markers flanking the  
1172chromosome V transmission ratio distortion peak and also homozygous *qgl/s5*.  
1173

1174Inheritance at Medea loci was tracked using PCR to amplify insertion/deletion markers near the  
1175transmission ratio distortion peaks:

1176LG3.1336F	TTAGAGCCGCTTGAAGTTGG
1177LG3.1336R	TCCGATGGACTAGGTTTCGT
1178LG5.2017F	TAACGCAATGGCCTCCTATC
1179LG5.2017R	GTTTGCTGGGTGGCCTAGTA

## 1180Simulations

1181We used simulations to investigate the effects of selfing on the spread of a maternal-toxin/zygotic  
1182antidote haplotype in populations with the distinctive andro dioecious mating system of *C. tropicalis*.  
1183Each simulated individual has a genotype and a sex. The single locus has haplotypes D and d. In  
1184single Medea simulations, D carries a maternal-effect toxin and zygotic-effect antidote, and d carries  
1185neither. In antagonistic Medea simulations, D and d carry two different Medea haplotypes. We initiate a  
1186population with *N* individuals, and genotype frequencies and sex ratio that are at equilibrium given the  
1187population's (fixed) selfing rate S and frequency *p* of the haplotype D. The neutral equilibrium  
1188inbreeding coefficient  $\hat{F}$  is  $S/(2-S)$  and the male frequency is  $(1-\hat{F})/2$ . During the simulations S is fixed  
1189but *F* can be far from its equilibrium value due to selection.  
1190

1191For simulations in Figure 10A and B, with fixed population sizes, we generated starting populations  
1192using the equilibrium frequencies below.

1193

1194Genotype	Frequency
1195DD herm	$((1-F)p^2 + pF)(1+F)/2$
1196Dd herm	$(1-F)2p(1-p)(1+F)/2$
1197dd herm	$((1-F)(1-p)^2 + (1-p)F)(1+F)/2$
1198DD male	$((1-F)p^2 + pF)(1-F)/2$
1199Dd male	$(1-F)2p(1-p)(1-F)/2$
1200dd male	$((1-F)(1-p)^2 + (1-p)F)(1-F)/2$

1201

1202For simulations in Figure 10C, assessing the effects of patch dynamics on frequencies of antagonistic  
1203Medeas, we draw genotype frequencies from a multinomial according to the equilibrium frequencies,  
1204but we assign sexes deterministically to ensure every patch receives hermaphrodite founders.

1205

1206We modeled the *C. tropicalis* androdioecious mating system, with self-fertile hermaphrodites that are  
1207incapable of mating with one another, and males that can cross-fertilize hermaphrodite oocytes. If no  
1208male is present in a population, each hermaphrodite produces a brood of size  $B$  by selfing. If there are  
1209males, each hermaphrodite produces  $SB$  hermaphrodite offspring by selfing and  $(1-S)B$  offspring by  
1210mating (male or hermaphrodite with equal probability), with a single father drawn randomly from the  
1211population of males. Self progeny are mostly hermaphrodites, except each has probability  $Him$  of being  
1212male ( $Him$  is the worm community name for the rate of male production by X-chromosome  
1213nondisjunction in selfing hermaphrodites).

1214

1215For simplicity, we assume that all selection is on embryo viability and that there is no cost to the Medea  
1216allele. For single-Medea analyses (Fig 10A and B), individuals that are dd but have Dd mothers are  
1217viable with probability  $V$  (i.e., if  $V>0$  some embryos can survive the maternal-effect toxin). Everybody  
1218else has viability 1.

1219

offspring			
Mother	DD	Dd	dd
DD	1	1	-
Dd	1	1	$V$
dd	-	1	1

1225

1226In simulations with antagonistic Medeas, we have two classes of affected offspring, with viabilities  $V_1$   
1227and  $V_2$ :

1228

offspring			
Mother	DD	Dd	dd
DD	1	1	-
Dd	$V_1$	1	$V_2$
dd	-	1	1

1234

1235In simulations with fixed population size, each discrete generation is sampled from the pool of viable  
1236embryos. We then track  $p$ , the frequency of the D allele, given parameters  $N$ ,  $S$ ,  $B$ ,  $V$ , and  $Him$ . In  
1237simulations with exponential growth,  $p$  and  $N$  are both variables. Simulations described in the text used  
1238 $B = 50$ ,  $V = V_1 = V_2 = 0.05$ , and  $Him = 0.005$ , except for the case of unequal penetrances of antagonistic  
1239Medea elements, where we set  $V_1$  and  $V_2$  to 0.4 and 0.05 to model chromosome V. We then investigate  
1240the effects of population size ( $N$ ) and selfing rate ( $S$ ) on Medea allele frequency  $p$ . Simulations started  
1241with allele frequency of 0.05 for single-Medea scenarios, and with frequency 0.2 for antagonistic-Medea  
1242scenarios. Simulation code is available on [github](#).

## 1243Source Data Files

1244Figure 1 - source data 1: outcrossProbability.tsv.zip; wild isolate outcross probability trials.

1245Figure 1 - source data 2: malePassaging.tsv.zip; wild isolate spontaneous male frequency.

1246Figure 2 - source data 1: selfer\_theta\_20kb.tsv.zip; Binned nucleotide diversity for *C. elegans*, *C.*  
1247 *briggsae*, and *C. tropicalis*.  
1248Figure 2 - source data 2: JU1373-NIC58.alignmentCoverage.tsv.zip; JU1373 and NIC58 identity and  
1249 copy number variation (Minimap2 alignment).  
1250Figure 2 - source data 3: theta\_10bp.bed.zip; *C. tropicalis* fine-scale nucleotide diversity (10 bp scale).  
1251Figure 3 - source data 1: isolateMetadata.tsv.zip; metadata for *C. tropicalis* wild isolates.  
1252Figure 4 - source data 1: RIL\_mating.tsv.zip; RIL outcross probability trials.  
1253Figure 5 - source data 1: RIL\_distortion.tsv.zip; genotype tables at transmission ratio distortion peaks  
1254 on chromosomes I, III and V.  
1255Figure 7 - source data 1: NIC58\_JU1373\_RIL\_crosses.tsv.zip; plate-level cross compatibility data for  
1256 JU1373, NIC58 and RILs.  
1257Figure 9 - source data 1: NIC58\_JU1373\_isolate\_crosses.tsv.zip; plate-level cross compatibility data for  
1258 JU1373, NIC58 and wild isolates.

## 1259Supplementary Files

1260Supplementary File 1: NIC58\_rqtICross.rda.zip; R/qtI cross object containing the NIC58 genetic map  
1261 and associated RIL genotypes.  
1262Supplementary File 2: tropicalisGenomes.zip; archive containing nuclear and mitochondrial genomes  
1263 and annotations for NIC58 and JU1373.  
1264Supplementary File 3: rawVariantCalls.zip; archive containing unfiltered variant calls for nuclear and  
1265 mitochondrial genomes.  
1266Supplementary File 4: filteredVariantCalls.zip; archive containing hard-filtered variant calls for the  
1267 nuclear genome.  
1268Supplementary File 5: processedVariantCalls.zip; archive containing hard-filtered variant calls for  
1269 nuclear and mitochondrial genomes with no missing data.  
1270Supplementary File 6: caeno\_orthogroups.tsv.zip; ortholog groupings for *Caenorhabditis* species.

## 1271Acknowledgements

1272We thank Arielle Martel, Jia Shen, and Patrick Ammerman for assistance in the lab, and the Félix,  
1273Teotónio, and Rockman labs for discussion. We are grateful to Marie-Anne Félix for the use of the JU  
1274strains, for stimulating discussions with her and Henrique Teotónio, and for helpful comments on the  
1275preprint from Asher Cutter. This work was supported by R01GM121828 (MVR), R01GM089972 (MVR),  
1276R01ES029930 (ECA and MVR), Dean's Undergraduate Research Fund grants (JY), the Centre  
1277National de la Recherche Scientifique (CNRS; CB), and the European Commission Marie Skłodowska-  
1278Curie Fellowship H2020-MSCA-IF-2017-798083 (LMN). *C. tropicalis* strain QG843 was collected under  
1279permit SEX/A-25-12 from the Republic of Panama. Sequencing data were generated by the Duke  
1280University Center for Genomic and Computational Biology and the New York University Center for  
1281Genomics and Systems Biology Core Facility, and this work was supported in part through the NYU IT  
1282High Performance Computing resources, services, and staff expertise.  
1283

## 1284References

1285 Andersen, E. C., Gerke, J. P., Shapiro, J. A., Crissman, J. R., Ghosh, R., Bloom, J. S., Félix, M.-A.,  
1286 & Kruglyak, L. (2012). Chromosome-scale selective sweeps shape *Caenorhabditis elegans*

- 1287 genomic diversity. *Nature Genetics*, 44(3), 285–290.
- 1288 Anderson, J. L., Morran, L. T., & Phillips, P. C. (2010). Outcrossing and the maintenance of males  
1289 within *C. elegans* populations. *The Journal of Heredity*, 101 Suppl 1, S62–S74.  
1290 <https://doi.org/10.1093/jhered/esq003>
- 1291 Andolfatto, P., Davison, D., Ereyilmaz, D., Hu, T. T., Mast, J., Sunayama-Morita, T., & Stern, D. L.  
1292 (2011). Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome  
1293 Research*, 21(4), 610–617. <https://doi.org/10.1101/gr.115402.110>
- 1294 Ayad, L. A. K., & Pissis, S. P. (2017). MARS: improving multiple circular sequence alignment using  
1295 refined sequences. *BMC Genomics*, 18(1), 1–10. <https://doi.org/10.1186/s12864-016-3477-5>
- 1296 Baird, S. E., & Stonesifer, R. (2012). Reproductive isolation in *Caenorhabditis briggsae*: Dysgenic  
1297 interactions between maternal- and zygotic-effect loci result in a delayed development  
1298 phenotype. *Worm*, 1(4), 189–195. <https://doi.org/10.4161/worm.23535>
- 1299 Baker, H. G. (1955). Self-Compatibility and Establishment After “Long-Distance” Dispersal.  
1300 *Evolution; International Journal of Organic Evolution*, 9(3), 347–349.  
1301 <https://doi.org/10.2307/2405656>
- 1302 Barrière, A., & Félix, M.-A. (2007). Temporal dynamics and linkage disequilibrium in natural  
1303 *Caenorhabditis elegans* populations. *Genetics*, 176(2), 999–1011.
- 1304 Barrière, A., Yang, S.-P., Pekarek, E., Thomas, C. G., Haag, E. S., & Ruvinsky, I. (2009). Detecting  
1305 heterozygosity in shotgun genome assemblies: Lessons from obligately outcrossing  
1306 nematodes. *Genome Research*, 19(3), 470–480. <https://doi.org/10.1101/gr.081851.108>
- 1307 Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using  
1308 lme4. *Journal of Statistical Software, Articles*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- 1309 Baym, M., Kryazhimskiy, S., Lieberman, T. D., Chung, H., Desai, M. M., & Kishony, R. (2015).  
1310 Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS One*, 10(5),  
1311 e0128036. <https://doi.org/10.1371/journal.pone.0128036>
- 1312 Beeman, R. W., & Friesen, K. S. (1999). Properties and natural occurrence of maternal-effect  
1313 selfish genes ('Medea' factors) in the red flour beetle, *Tribolium castaneum*. *Heredity*, 82 (Pt 5),

- 1314 529–534. <https://doi.org/10.1038/sj.hdy.6885150>
- 1315 Beeman, R. W., Friesen, K. S., & Denell, R. E. (1992). Maternal-effect selfish genes in flour beetles.
- 1316 *Science*, 256(5053), 89–92. <https://doi.org/10.1126/science.1566060>
- 1317 Ben-David, E., Burga, A., & Kruglyak, L. (2017). A maternal-effect selfish genetic element in
- 1318 *Caenorhabditis elegans*. *Science*, 356(6342), 1051–1055.
- 1319 <https://doi.org/10.1126/science.aan0621>
- 1320 Ben-David, E., Pliota, P., Widen, S. A., Koreshova, A., Lemus-Vergara, T., Verpukhovskiy, P.,
- 1321 Mandali, S., Braendle, C., Burga, A., & Kruglyak, L. (2020). *Ubiquitous selfish toxin-antidote*
- 1322 *elements in Caenorhabditis species*. <https://doi.org/10.1101/2020.08.06.240564>
- 1323 Bernstein, M. R., & Rockman, M. V. (2016). Fine-Scale crossover rate variation on the
- 1324 *Caenorhabditis elegans* X chromosome. *G3: Genes, Genomes, Genetics*, 6(6), 1767–1776.
- 1325 Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf,
- 1326 M., & Stadler, P. F. (2013). MITOS: improved de novo metazoan mitochondrial genome
- 1327 annotation. *Molecular Phylogenetics and Evolution*, 69(2), 313–319.
- 1328 <https://doi.org/10.1016/j.ympev.2012.08.023>
- 1329 Brand, C. L., Larracuente, A. M., & Presgraves, D. C. (2015). Origin, evolution, and population
- 1330 genetics of the selfish Segregation Distorter gene duplication in European and African
- 1331 populations of *Drosophila melanogaster*. *Evolution; International Journal of Organic Evolution*,
- 1332 69(5), 1271–1283. <https://doi.org/10.1111/evo.12658>
- 1333 Brenner, S. (1974). The Genetics of *Caenorhabditis elegans*. *Genetics*, 77(1), 71–94.
- 1334 <https://www.genetics.org/content/77/1/71>
- 1335 Broman, K. W., Wu, H., Sen, S., & Churchill, G. A. (2003). R/qtl: QTL mapping in experimental
- 1336 crosses. *Bioinformatics*, 19(7), 889–890. <https://doi.org/10.1093/bioinformatics/btg112>
- 1337 Bull, J. J. (2017). Lethal gene drive selects inbreeding. *Evolution, Medicine, and Public Health*,
- 1338 2017(1), 1–16. <https://doi.org/10.1093/emph/eow030>
- 1339 Bull, J. J., Remien, C. H., & Krone, S. M. (2019). Gene-drive-mediated extinction is thwarted by
- 1340 population structure and evolution of sib mating. *Evolution, Medicine, and Public Health*,

- 1341 2019(1), 66–81. <https://doi.org/10.1093/emph/eoz014>
- 1342 Bulmer, M. G., & Taylor, P. D. (1980). Sex ratio under the haystack model. *Journal of Theoretical*  
1343 *Biology*, 86(1), 83–89. [https://doi.org/10.1016/0022-5193\(80\)90066-1](https://doi.org/10.1016/0022-5193(80)90066-1)
- 1344 Burt, A. (2003). Site-specific selfish genes as tools for the control and genetic engineering of natural  
1345 populations. *Proceedings. Biological Sciences / The Royal Society*, 270(1518), 921–928.  
1346 <https://doi.org/10.1098/rspb.2002.2319>
- 1347 Champer, J., Liu, J., Oh, S. Y., Reeves, R., Luthra, A., Oakes, N., Clark, A. G., & Messer, P. W.  
1348 (2018). Reducing resistance allele formation in CRISPR gene drive. *Proceedings of the*  
1349 *National Academy of Sciences of the United States of America*, 115(21), 5522–5527.  
1350 <https://doi.org/10.1073/pnas.1720354115>
- 1351 Charlesworth, B. (2012). The effects of deleterious mutations on evolution at linked sites. *Genetics*,  
1352 190(1), 5–22. <https://doi.org/10.1534/genetics.111.134288>
- 1353 Charlesworth, B., Nordborg, M., & Charlesworth, D. (1997). The effects of local selection, balanced  
1354 polymorphism and background selection on equilibrium patterns of genetic diversity in  
1355 subdivided populations. *Genetical Research*, 70(2), 155–174.  
1356 <https://doi.org/10.1017/s0016672397002954>
- 1357 Chaudhuri, J., Bose, N., Tandonnet, S., Adams, S., Zuco, G., Kache, V., Parihar, M., von Reuss, S.  
1358 H., Schroeder, F. C., & Pires-daSilva, A. (2015). Mating dynamics in a nematode with three  
1359 sexes and its evolutionary implications. *Scientific Reports*, 5(1), 1–11.  
1360 <https://doi.org/10.1038/srep17676>
- 1361 Chelo, I. M., Afonso, B., Carvalho, S., Theologidis, I., Goy, C., Pino-Querido, A., Proulx, S., &  
1362 Teotónio, H. (2019). Partial selfing can reduce genetic loads while maintaining diversity during  
1363 evolution. *G3*, 9, 2811–2821. <https://www.g3journal.org/content/9/9/2811>
- 1364 Courret, C., Chang, C.-H., Wei, K. H.-C., Montchamp-Moreau, C., & Larracuente, A. M. (2019).  
1365 Meiotic drive mechanisms: lessons from *Drosophila*. *Proceedings. Biological Sciences / The*  
1366 *Royal Society*, 286(1913), 20191430. <https://doi.org/10.1098/rspb.2019.1430>
- 1367 Crombie, T. A., Zdraljevic, S., Cook, D. E., Tanny, R. E., Brady, S. C., Wang, Y., Evans, K. S.,

- 1368 Hahnel, S., Lee, D., Rodriguez, B. C., Zhang, G., van der Zwagg, J., Kiontke, K., & Andersen,  
1369 E. C. (2019). Deep sampling of Hawaiian *Caenorhabditis elegans* reveals high genetic diversity  
1370 and admixture with global populations. *eLife*, 8, e50465. <https://doi.org/10.7554/eLife.50465>
- 1371 Cutter, A. D. (2015). *Caenorhabditis* evolution in the wild. *BioEssays: News and Reviews in*  
1372 *Molecular, Cellular and Developmental Biology*, 37(9), 983–995.  
1373 <https://doi.org/10.1002/bies.201500053>
- 1374 Cutter, A. D. (2019). Reproductive transitions in plants and animals: selfing syndrome, sexual  
1375 selection and speciation. *The New Phytologist*, 224(3), 1080–1094.  
1376 <https://doi.org/10.1111/nph.16075>
- 1377 Cutter, A. D., Félix, M.-A., Barrière, A., & Charlesworth, D. (2006). Patterns of nucleotide  
1378 polymorphism distinguish temperate and tropical wild isolates of *Caenorhabditis briggsae*.  
1379 *Genetics*, 173(4), 2021–2031. <https://doi.org/10.1534/genetics.106.058651>
- 1380 DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A.,  
1381 del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernytsky, A. M.,  
1382 Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., & Daly, M. J. (2011). A framework  
1383 for variation discovery and genotyping using next-generation DNA sequencing data. *Nature*  
1384 *Genetics*, 43(5), 491–498. <https://doi.org/10.1038/ng.806>
- 1385 Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., &  
1386 Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), 15–  
1387 21. <https://doi.org/10.1093/bioinformatics/bts635>
- 1388 Dolgin, E. S., Charlesworth, B., Baird, S. E., & Cutter, A. D. (2007). Inbreeding and Outbreeding  
1389 Depression in *Caenorhabditis* Nematodes. *Evolution; International Journal of Organic*  
1390 *Evolution*, 61(6), 1339–1352. <https://www.jstor.org/stable/4621380>
- 1391 Dolgin, E. S., Félix, M.-A., & Cutter, A. D. (2008). Hakuna Nematoda: genetic and phenotypic  
1392 diversity in African isolates of *Caenorhabditis elegans*<i> and *C. briggsae*. *Heredity*, 100(3),  
1393 304–315. <https://doi.org/10.1038/sj.hdy.6801079>
- 1394 Dowle, M., & Srinivasan, A. (2019). *data.table: Extension of ‘data.frame’*. <https://CRAN.R-project.org/package=data.table>

- 1395 project.org/package=data.table
- 1396 Drown, D. M., & Wade, M. J. (2014). Runaway coevolution: adaptation to heritable and nonheritable  
1397 environments. *Evolution; International Journal of Organic Evolution*, 68(10), 3039–3046.
- 1398 <https://doi.org/10.1111/evo.12470>
- 1399 Drury, D. W., Dapper, A. L., Siniard, D. J., Zentner, G. E., & Wade, M. J. (2017). CRISPR/Cas9  
1400 gene drives in genetically variable and nonrandomly mating wild populations. *Science  
1401 Advances*, 3(5), e1601910. <https://doi.org/10.1126/sciadv.1601910>
- 1402 Dunn, P. K., & Smyth, G. K. (2016). *dglm: Double Generalized Linear Models*. <https://CRAN.R-project.org/package=dglm>
- 1403 Duong, T. (2020). *ks: Kernel Smoothing*. <https://CRAN.R-project.org/package=ks>
- 1404 Ellis, R. E. (2017). “The persistence of memory”—Hermaphroditism in nematodes. *Molecular  
1405 Reproduction and Development*, 84(2), 144–157. <https://doi.org/10.1002/mrd.22668>
- 1406 Escobar, J. S., Auld, J. R., Correa, A. C., Alonso, J. M., Bony, Y. K., Coutellec, M.-A., Koene, J. M.,  
1407 Pointier, J.-P., Jarne, P., & David, P. (2011). PATTERNS OF MATING-SYSTEM EVOLUTION  
1408 IN HERMAPHRODITIC ANIMALS: CORRELATIONS AMONG SELFING RATE, INBREEDING  
1409 DEPRESSION, AND THE TIMING OF REPRODUCTION. *Evolution; International Journal of  
1410 Organic Evolution*, 65(5), 1233–1253. <https://doi.org/10.1111/j.1558-5646.2011.01218.x>
- 1411 Evans, T. C., & Hunter, C. P. (2005). Translational control of maternal RNAs. In The *C. elegans*  
1412 Research Community (Ed.), *WormBook*. <https://doi.org/doi/10.1895/wormbook.1.34.1>
- 1413 Félix, M.-A. (2020, August). *Félix database*. World Wide Worms.  
1414 <https://justbio.com/worldwideworms/>
- 1415 Félix, M.-A., & Braendle, C. (2010). The natural history of *Caenorhabditis elegans*. *Current Biology:  
1416 CB*, 20(22), R965–R969. <https://doi.org/10.1016/j.cub.2010.09.050>
- 1417 Félix, M.-A., Braendle, C., & Cutter, A. D. (2014). A Streamlined System for Species Diagnosis in  
1418 Caenorhabditis (Nematoda: Rhabditidae) with Name Designations for 15 Distinct Biological  
1419 Species. *PloS One*, 9(4), e94723. <https://doi.org/10.1371/journal.pone.0094723>
- 1420 Félix, M.-A., & Duveau, F. (2012). Population dynamics and habitat sharing of natural populations of

- 1422      *Caenorhabditis elegans* and *C. briggsae*. *BMC Biology*, 10, 59. <https://doi.org/10.1186/1741-7007-10-59>
- 1423      Félix, M.-A., Jovelin, R., Ferrari, C., Han, S., Cho, Y. R., Andersen, E. C., Cutter, A. D., & Braendle, C. (2013). Species richness, distribution and genetic diversity of *Caenorhabditis* nematodes in a remote tropical rainforest. *BMC Evolutionary Biology*, 13(1), 10. <https://doi.org/10.1186/1471-2148-13-10>
- 1424      Ferrari, C., Salle, R., Callemeyn-Torre, N., Jovelin, R., Cutter, A. D., & Braendle, C. (2017).  
1425      Ephemeral-habitat colonization and neotropical species richness of *Caenorhabditis* nematodes.  
1426      *BMC Ecology*, 17(1), 43. <https://doi.org/10.1186/s12898-017-0150-z>
- 1427      Fierst, J. L., Willis, J. H., Thomas, C. G., Wang, W., Reynolds, R. M., Ahearne, T. E., Cutter, A. D., & Phillips, P. C. (2015). Reproductive Mode and the Evolution of Genome Size and Structure in *Caenorhabditis* Nematodes. *PLoS Genetics*, 11(6), e1005323.  
1428      <https://doi.org/10.1371/journal.pgen.1005323>
- 1429      Fishman, L., & Sweigart, A. L. (2018). When Two Rights Make a Wrong: The Evolutionary Genetics  
1430      of Plant Hybrid Incompatibilities. *Annual Review of Plant Biology*, 69, 707–731.  
1431      <https://doi.org/10.1146/annurev-arplant-042817-040113>
- 1432      Fradin, H., Kiontke, K., Zegar, C., Gutwein, M., Lucas, J., Kovtun, M., Corcoran, D. L., Baugh, L. R.,  
1433      Fitch, D. H. A., Piano, F., & Gunsalus, K. C. (2017). Genome Architecture and Evolution of a  
1434      Unichromosomal Asexual Nematode. *Current Biology: CB*, 27(19), 2928–2939.e6.  
1435      <https://doi.org/10.1016/j.cub.2017.08.038>
- 1436      Frøkjaer-Jensen, C., Davis, M. W., Hopkins, C. E., Newman, B. J., Thummel, J. M., Olesen, S.-P.,  
1437      Grunnet, M., & Jorgensen, E. M. (2008). Single-copy insertion of transgenes in *Caenorhabditis*  
1438      *elegans*. *Nature Genetics*, 40(11), 1375–1383. <https://doi.org/10.1038/ng.248>
- 1439      Garrison, E., Sirén, J., Novak, A. M., Hickey, G., Eizenga, J. M., Dawson, E. T., Jones, W., Garg,  
1440      S., Markello, C., Lin, M. F., Paten, B., & Durbin, R. (2018). Variation graph toolkit improves  
1441      read mapping by representing genetic variation in the reference. *Nature Biotechnology*, 36(9),  
1442      875–879. <https://doi.org/10.1038/nbt.4227>

- 1449 Gimond, C., Jovelin, R., Han, S., Ferrari, C., Cutter, A. D., & Braendle, C. (2013). Outbreeding  
1450 Depression with Low Genetic Variation in Selfing *Caenorhabditis* Nematodes. *Evolution; International Journal of Organic Evolution*, 67(11), 3087–3101.  
1451  
1452 <https://doi.org/10.1111/evo.12203>
- 1453 Goodwillie, C., Kalisz, S., & Eckert, C. G. (2005). The Evolutionary Enigma of Mixed Mating  
1454 Systems in Plants: Occurrence, Theoretical Explanations, and Empirical Evidence. *Annual Review of Ecology, Evolution, and Systematics*, 36(1), 47–79.  
1455  
1456 <https://doi.org/10.1146/annurev.ecolsys.36.091704.175539>
- 1457 Grosmaire, M., Launay, C., Siegwald, M., Brugiére, T., Estrada-Virrueta, L., Berger, D., Burny, C.,  
1458 Modolo, L., Blaxter, M., Meister, P., Félix, M.-A., Gouyon, P.-H., & Delattre, M. (2019). Males  
1459 as somatic investment in a parthenogenetic nematode. *Science*, 363(6432), 1210–1213.  
1460  
<https://doi.org/10.1126/science.aau0099>
- 1461 Grundt, H. H., Kjølner, S., Borgen, L., Rieseberg, L. H., & Brochmann, C. (2006). High biological  
1462 species diversity in the arctic flora. *Proceedings of the National Academy of Sciences of the United States of America*, 103(4), 972–975. <https://doi.org/10.1073/pnas.0510270103>  
1463
- 1464 Haber, M., Schüngel, M., Putz, A., Müller, S., Hasert, B., & Schulenburg, H. (2005). Evolutionary  
1465 history of *Caenorhabditis elegans* inferred from microsatellites: evidence for spatial and  
1466 temporal genetic differentiation and the occurrence of outbreeding. *Molecular Biology and Evolution*, 22(1), 160–173. <https://doi.org/10.1093/molbev/msh264>  
1467
- 1468 Hedrick, P. W. (2013). Adaptive introgression in animals: examples and comparison to new  
1469 mutation and standing variation as sources of adaptive variation. *Molecular Ecology*, 22(18),  
1470 4606–4618. <https://doi.org/10.1111/mec.12415>
- 1471 Hoff, K. J., Lomsadze, A., Borodovsky, M., & Stanke, M. (2019). Whole-Genome Annotation with  
1472 BRAKER. In M. Kollmar (Ed.), *Gene Prediction: Methods and Protocols* (pp. 65–95). Springer  
1473 New York. [https://doi.org/10.1007/978-1-4939-9173-0\\_5](https://doi.org/10.1007/978-1-4939-9173-0_5)
- 1474 Hubley, R., Finn, R. D., Clements, J., Eddy, S. R., Jones, T. A., Bao, W., Smit, A. F. A., & Wheeler,  
1475 T. J. (2016). The Dfam database of repetitive DNA families. *Nucleic Acids Research*, 44(D1),

- 1476 D81–D89. <https://doi.org/10.1093/nar/gkv1272>
- 1477 Igic, B., & Kohn, J. R. (2006). The Distribution of Plant Mating Systems: Study Bias Against  
1478 Obligately Outcrossing Species. *Evolution; International Journal of Organic Evolution*, 60(5),  
1479 1098–1103. <https://doi.org/10.1111/j.0014-3820.2006.tb01186.x>
- 1480 Jalinsky, J., Logsdon, J. M., Jr, & Neiman, M. (2020). Male phenotypes in a female framework:  
1481 Evidence for degeneration in sperm produced by male snails from asexual lineages. *Journal of  
1482 Evolutionary Biology*. <https://doi.org/10.1111/jeb.13632>
- 1483 Jarne, P., & Auld, J. R. (2006). Animals Mix It up Too: The Distribution of Self-Fertilization Among  
1484 Hermaphroditic Animals. *Evolution; International Journal of Organic Evolution*, 60(9), 1816–  
1485 1824. <https://doi.org/10.1111/j.0014-3820.2006.tb00525.x>
- 1486 Kahle, D., & Wickham, H. (2013). ggmap: Spatial Visualization with ggplot2. In *The R Journal* (Vol.  
1487 5, Issue 1, pp. 144–161). <https://journal.r-project.org/archive/2013-1/kahle-wickham.pdf>
- 1488 Kajitani, R., Toshimoto, K., Noguchi, H., Toyoda, A., Ogura, Y., Okuno, M., Yabana, M., Harada,  
1489 M., Nagayasu, E., Maruyama, H., Kohara, Y., Fujiyama, A., Hayashi, T., & Itoh, T. (2014).  
1490 Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short  
1491 reads. *Genome Research*, 24(8), 1384–1395. <https://doi.org/10.1101/gr.170720.113>
- 1492 Kamran-Disfani, A., & Agrawal, A. F. (2014). Selfing, adaptation and background selection in finite  
1493 populations. *Journal of Evolutionary Biology*, 27(7), 1360–1371.
- 1494 Kanzaki, N., Kiontke, K., Tanaka, R., Hirooka, Y., Schwarz, A., Müller-Reichert, T., Chaudhuri, J., &  
1495 Pires-daSilva, A. (2017). Description of two three-gendered nematode species in the new  
1496 genus *Auanema* (Rhabditina) that are models for reproductive mode evolution. *Scientific  
1497 Reports*, 7(1), 11135. <https://doi.org/10.1038/s41598-017-09871-1>
- 1498 Kaur, T., & Rockman, M. V. (2014). Crossover heterogeneity in the absence of hotspots in  
1499 *Caenorhabditis elegans*. *Genetics*, 196(1), 137–148.
- 1500 Kiontke, K. C. (2005). The phylogenetic relationships of *Caenorhabditis* and other rhabditids. In The  
1501 *C. elegans* Research Community (Ed.), *WormBook*. <https://doi.org/10.1895/wormbook.1.11.1>
- 1502 Kiontke, K. C., Félix, M.-A., Ailion, M., Rockman, M. V., Braendle, C., Pénigault, J.-B., & Fitch, D. H.

- 1503 A. (2011). A phylogeny and molecular barcodes for *Caenorhabditis*, with numerous new  
1504 species from rotting fruits. *BMC Evolutionary Biology*, 11(1), 339. <https://doi.org/10.1186/1471->  
1505 2148-11-339
- 1506 Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads  
1507 using repeat graphs. *Nature Biotechnology*, 37(5), 540–546. <https://doi.org/10.1038/s41587->  
1508 019-0072-8
- 1509 Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu:  
1510 scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation.  
1511 *Genome Research*, 27(5), 722–736. <https://doi.org/10.1101/gr.215087.116>
- 1512 Köster, J., & Rahmann, S. (2012). Snakemake--a scalable bioinformatics workflow engine.  
1513 *Bioinformatics* , 28(19), 2520–2522. <https://doi.org/10.1093/bioinformatics/bts480>
- 1514 Lande, R., & Schemske, D. W. (1985). THE EVOLUTION OF SELF-FERTILIZATION AND  
1515 INBREEDING DEPRESSION IN PLANTS. I. GENETIC MODELS. *Evolution; International*  
1516 *Journal of Organic Evolution*, 39(1), 24–40. <https://doi.org/10.1111/j.1558-5646.1985.tb04077.x>
- 1517 Lee, D., Zdraljevic, S., Stevens, L., Wang, Y., Tanny, R. E., Crombie, T. A., Cook, D. E., Webster,  
1518 A. K., Chirakar, R., Ryan Baugh, L., Sterken, M. G., Braendle, C., Félix, M.-A., Rockman, M. V.,  
1519 & Andersen, E. C. (2020). *Balancing selection maintains ancient genetic diversity in C. elegans*  
1520 (p. 2020.07.23.218420). <https://doi.org/10.1101/2020.07.23.218420>
- 1521 Lemire, B. (2005). Mitochondrial genetics. In The *C. elegans* Research Community (Ed.),  
1522 *WormBook*. <https://doi.org/10.1895/wormbook.1.25.1>
- 1523 Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and  
1524 population genetical parameter estimation from sequencing data. *Bioinformatics* , 27(21),  
1525 2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>
- 1526 Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* , 34(18),  
1527 3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>
- 1528 Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler  
1529 transform. *Bioinformatics* , 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>

- 1530 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G.,  
1531 Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence  
1532 Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079.  
1533 <https://doi.org/10.1093/bioinformatics/btp352>
- 1534 Lindholm, A. K., Dyer, K. A., Firman, R. C., Fishman, L., Forstmeier, W., Holman, L., Johannesson,  
1535 H., Knief, U., Kokko, H., Larracuente, A. M., Manser, A., Montchamp-Moreau, C., Petrosyan, V.  
1536 G., Pomiankowski, A., Presgraves, D. C., Safranova, L. D., Sutter, A., Unckless, R. L.,  
1537 Verspoor, R. L., ... Price, T. A. R. (2016). The Ecology and Evolutionary Dynamics of Meiotic  
1538 Drive. *Trends in Ecology & Evolution*, 31(4), 315–326.  
1539 <https://doi.org/10.1016/j.tree.2016.02.001>
- 1540 Li, R., Ren, X., Bi, Y., Ding, Q., Ho, V. W. S., & Zhao, Z. (2018). Comparative mitochondrial  
1541 genomics reveals a possible role of a recent duplication of NADH dehydrogenase subunit 5 in  
1542 gene regulation. *DNA Research: An International Journal for Rapid Publication of Reports on*  
1543 *Genes and Genomes*, 25(6), 577–586. <https://doi.org/10.1093/dnares/dsy026>
- 1544 Lively, C. M., & Lloyd, D. G. (1990). The Cost of Biparental Sex Under Individual Selection. *The*  
1545 *American Naturalist*, 135(4), 489–500. <https://doi.org/10.1086/285058>
- 1546 Lorenzen, M. D., Gnirke, A., Margolis, J., Garnes, J., Campbell, M., Stuart, J. J., Aggarwal, R.,  
1547 Richards, S., Park, Y., & Beeman, R. W. (2008). The maternal-effect, selfish genetic element  
1548 Medea is associated with a composite Tc1 transposon. *Proceedings of the National Academy*  
1549 *of Sciences of the United States of America*, 105(29), 10085–10089.  
1550 <https://doi.org/10.1073/pnas.0800444105>
- 1551 Lyttle, T. W. (1991). Segregation distorters. *Annual Review of Genetics*, 25, 511–557.  
1552 <https://doi.org/10.1146/annurev.ge.25.120191.002455>
- 1553 Maheshwari, S., & Barbash, D. A. (2011). The Genetics of Hybrid Incompatibilities. *Annual Review*  
1554 *of Genetics*, 45(1), 331–355. <https://doi.org/10.1146/annurev-genet-110410-132514>
- 1555 Mayer, W. E., Herrmann, M., & Sommer, R. J. (2007). Phylogeny of the nematode genus  
1556 *Pristionchus* and implications for biodiversity, biogeography and the evolution of

- 1557 hermaphroditism. *BMC Evolutionary Biology*, 7(1), 1–13. <https://doi.org/10.1186/1471-2148-7-104>
- 1559 McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K.,  
1560 Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: a  
1561 MapReduce framework for analyzing next-generation DNA sequencing data. *Genome  
Research*, 20(9), 1297–1303.
- 1563 Morran, L. T., Parmenter, M. D., & Phillips, P. C. (2009). Mutation load and rapid adaptation favour  
1564 outcrossing over self-fertilization. *Nature*, 462(7271), 350–352.
- 1565 Morran, L. T., Schmidt, O. G., Gelarden, I. A., Parrish, R. C., 2nd, & Lively, C. M. (2011). Running  
1566 with the Red Queen: host-parasite coevolution selects for biparental sex. *Science*, 333(6039),  
1567 216–218. <https://doi.org/10.1126/science.1206360>
- 1568 Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction  
1569 endonucleases. *Proceedings of the National Academy of Sciences of the United States of  
1570 America*, 76(10), 5269–5273. <https://doi.org/10.1073/pnas.76.10.5269>
- 1571 Nigon, V., & Félix, M.-A. (2017). History of research on *C. elegans* and other free-living nematodes  
1572 as model organisms. In The *C. elegans* Research Community (Ed.), *WormBook*.  
1573 <https://doi.org/10.1895/wormbook.1.181.1>
- 1574 Noble, L. M., Chang, A. S., McNelis, D., Kramer, M., Yen, M., Nicodemus, J. P., Riccardi, D. D.,  
1575 Ammerman, P., Phillips, M., Islam, T., & Rockman, M. V. (2015). Natural Variation in plep-1  
1576 Causes Male-Male Copulatory Behavior in *C. elegans*. *Current Biology: CB*, 25(20), 2730–  
1577 2737. <https://doi.org/10.1016/j.cub.2015.09.019>
- 1578 Nordborg, M., Charlesworth, B., & Charlesworth, D. (1996). Increased levels of polymorphism  
1579 surrounding selectively maintained sites in highly selfing species. *Proceedings of the Royal  
1580 Society of London. Series B: Biological Sciences*, 263(1373), 1033–1039.  
1581 <https://doi.org/10.1098/rspb.1996.0152>
- 1582 Nordborg, M., & Donnelly, P. (1997). The Coalescent Process With Selfing. *Genetics*, 146(3),  
1583 1185–1195. <https://www.genetics.org/content/146/3/1185>

- 1584 Otto, S. P. (2009). The Evolutionary Enigma of Sex. *The American Naturalist*, 174(S1), S1–S14.
- 1585 <https://doi.org/10.1086/599084>
- 1586 Palopoli, M. F., Rockman, M. V., TinMaung, A., Ramsay, C., Curwen, S., Aduna, A., Laurita, J., &
- 1587 Kruglyak, L. (2008). Molecular basis of the copulatory plug polymorphism in *Caenorhabditis*
- 1588 *elegans*. *Nature*, 454(7207), 1019–1022.
- 1589 Peters, L. L., & Barker, J. E. (1993). Novel inheritance of the murine severe combined anemia and
- 1590 thrombocytopenia (Scat) phenotype. *Cell*, 74(1), 135–142. [https://doi.org/10.1016/0092-8674\(93\)90301-6](https://doi.org/10.1016/0092-8674(93)90301-6)
- 1592 Prasad, A., Croydon-Sugarman, M. J. F., Murray, R. L., & Cutter, A. D. (2011). Temperature-
- 1593 Dependent Fecundity Associates with Latitude in *Caenorhabditis Briggsae*. *Evolution; International Journal of Organic Evolution*, 65(1), 52–63. <https://doi.org/10.1111/j.1558-5646.2010.01110.x>
- 1596 Presgraves, D. C. (2010). The molecular evolutionary basis of species formation. *Nature Reviews. Genetics*, 11(3), 175–180. <https://doi.org/10.1038/nrg2718>
- 1598 Presgraves, D. C., Gérard, P. R., Cherukuri, A., & Lyttle, T. W. (2009). Large-scale selective sweep
- 1599 among Segregation Distorter chromosomes in African populations of *Drosophila melanogaster*.
- 1600 *PLoS Genetics*, 5(5), e1000463. <https://doi.org/10.1371/journal.pgen.1000463>
- 1601 Price, T. A. R., Verspoor, R., & Wedell, N. (2019). Ancient gene drives: an evolutionary paradox.
- 1602 *Proceedings. Biological Sciences / The Royal Society*, 286(1917), 20192267.
- 1603 <https://doi.org/10.1098/rspb.2019.2267>
- 1604 Price, T. A. R., Windbichler, N., Unckless, R. L., Sutter, A., Runge, J.-N., Ross, P. A.,
- 1605 Pomiankowski, A., Nuckolls, N. L., Montchamp-Moreau, C., Mideo, N., Martin, O. Y., Manser,
- 1606 A., Legros, M., Larracuente, A. M., Holman, L., Godwin, J., Gemmell, N., Courret, C.,
- 1607 Buchman, A., ... Lindholm, A. K. (2020). Resistance to natural and synthetic gene drive
- 1608 systems. *Journal of Evolutionary Biology*, 33(10), 1345–1360. <https://doi.org/10.1111/jeb.13693>
- 1609 R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for
- 1610 Statistical Computing. <https://www.R-project.org/>

- 1611 Richaud, A., Zhang, G., Lee, D., Lee, J., & Félix, M.-A. (2018). The Local Coexistence Pattern of  
1612 Selfing Genotypes in *Caenorhabditis elegans* Natural Metapopulations. *Genetics*, 208(2), 807–  
1613 821. <https://doi.org/10.1534/genetics.117.300564>
- 1614 Robertson, S., & Lin, R. (2015). The Maternal-to-Zygotic Transition in *C. elegans*. In H. D. Lipshitz  
1615 (Ed.), *Current Topics in Developmental Biology* (Vol. 113, pp. 1–42). Academic Press.  
1616 <https://doi.org/10.1016/bs.ctdb.2015.06.001>
- 1617 Ross, J. A., Koboldt, D. C., Staisch, J. E., Chamberlin, H. M., Gupta, B. P., Miller, R. D., Baird, S.  
1618 E., & Haag, E. S. (2011). *Caenorhabditis briggsae* Recombinant Inbred Line Genotypes Reveal  
1619 Inter-Strain Incompatibility and the Evolution of Recombination. *PLoS Genetics*, 7(7),  
1620 e1002174. <https://doi.org/10.1371/journal.pgen.1002174>
- 1621 Ruan, J., & Li, H. (2019). *Fast and accurate long-read assembly with wtdbg2* (p. 530972).  
1622 <https://doi.org/10.1101/530972>
- 1623 Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30(14), 2068–  
1624 2069. <https://doi.org/10.1093/bioinformatics/btu153>
- 1625 Seidel, H. S., Ailion, M., Li, J., van Oudenaarden, A., Rockman, M. V., & Kruglyak, L. (2011). A  
1626 Novel Sperm-Delivered Toxin Causes Late-Stage Embryo Lethality and Transmission Ratio  
1627 Distortion in *C. elegans*. *PLoS Biology*, 9(7), e1001115.  
1628 <https://doi.org/10.1371/journal.pbio.1001115>
- 1629 Seidel, H. S., Rockman, M. V., & Kruglyak, L. (2008). Widespread genetic incompatibility in *C.*  
1630 *elegans* maintained by balancing selection. *Science*, 319(5863), 589–594.  
1631 <https://doi.org/10.1126/science.1151107>
- 1632 Seppey, M., Manni, M., & Zdobnov, E. M. (2019). BUSCO: Assessing Genome Assembly and  
1633 Annotation Completeness. In M. Kollmar (Ed.), *Gene Prediction: Methods and Protocols* (pp.  
1634 227–245). Springer New York. [https://doi.org/10.1007/978-1-4939-9173-0\\_14](https://doi.org/10.1007/978-1-4939-9173-0_14)
- 1635 Shi, C., Runnels, A. M., & Murphy, C. T. (2017). Mating and male pheromone kill *Caenorhabditis*  
1636 males through distinct mechanisms. *eLife*, 6. <https://doi.org/10.7554/eLife.23493>
- 1637 Sivasundar, A., & Hey, J. (2005). Sampling from natural populations with RNAi reveals high

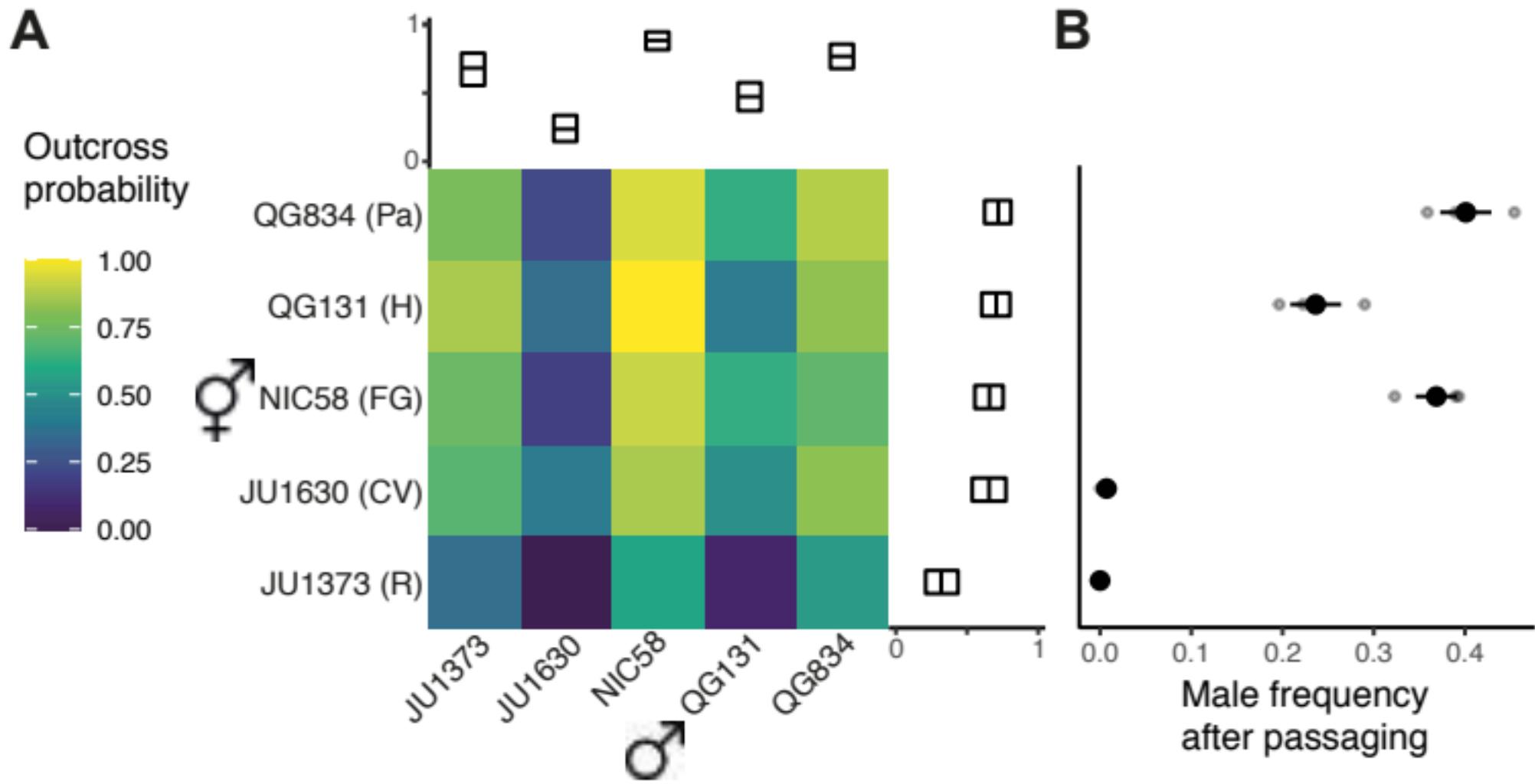
- 1638 outcrossing and population structure in *Caenorhabditis elegans*. *Current Biology: CB*, 15(17),  
1639 1598–1602. <https://doi.org/10.1016/j.cub.2005.08.034>
- 1640 Slowikowski, K. (2020). *ggrepel: Automatically Position Non-Overlapping Text Labels with “ggplot2.”*  
1641 <https://CRAN.R-project.org/package=ggrepel>
- 1642 Smit, A., Hubley, R., & Green, P. (n.d.). *RepeatMasker*. Retrieved August 6, 2020, from  
1643 <http://www.repeatmasker.org/>
- 1644 Stanke, M., Schöffmann, O., Morgenstern, B., & Waack, S. (2006). Gene prediction in eukaryotes  
1645 with a generalized hidden Markov model that uses hints from external sources. *BMC  
1646 Bioinformatics*, 7, 62. <https://doi.org/10.1186/1471-2105-7-62>
- 1647 Stewart, A. D., & Phillips, P. C. (2002). Selection and maintenance of androdioecy in  
1648 *Caenorhabditis elegans*. *Genetics*, 160(3), 975–982.
- 1649 Stiernagle, T. (2006). Maintenance of *C. elegans*. In The *C. elegans* Research Community (Ed.),  
1650 *WormBook*. <https://doi.org/10.1895/wormbook.1.101.1>
- 1651 Sunnucks, P., & Hales, D. F. (1996). Numerous transposed sequences of mitochondrial cytochrome  
1652 oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). *Molecular Biology and  
1653 Evolution*, 13(3), 510–524. <https://doi.org/10.1093/oxfordjournals.molbev.a025612>
- 1654 Sweigart, A. L., Brandvain, Y., & Fishman, L. (2019). Making a Murderer: The Evolutionary Framing  
1655 of Hybrid Gamete-Killers. *Trends in Genetics: TIG*, 35(4), 245–252.  
1656 <https://doi.org/10.1016/j.tig.2019.01.004>
- 1657 Teotonio, H., Carvalho, S., Manoel, D., Roque, M., & Chelo, I. M. (2012). Evolution of Outcrossing  
1658 in Experimental Populations of *Caenorhabditis elegans*. *PloS One*, 7(4), e35811.  
1659 <https://doi.org/10.1371/journal.pone.0035811>
- 1660 Teotónio, H., Manoel, D., & Phillips, P. C. (2006). Genetic variation for outcrossing among  
1661 *Caenorhabditis elegans* isolates. *Evolution; International Journal of Organic Evolution*, 60(6),  
1662 1300–1305.
- 1663 Theologidis, I., Chelo, I. M., Goy, C., & Teotónio, H. (2014). Reproductive assurance drives  
1664 transitions to self-fertilization in experimental *Caenorhabditis elegans*. *BMC Biology*, 12(1), 1–

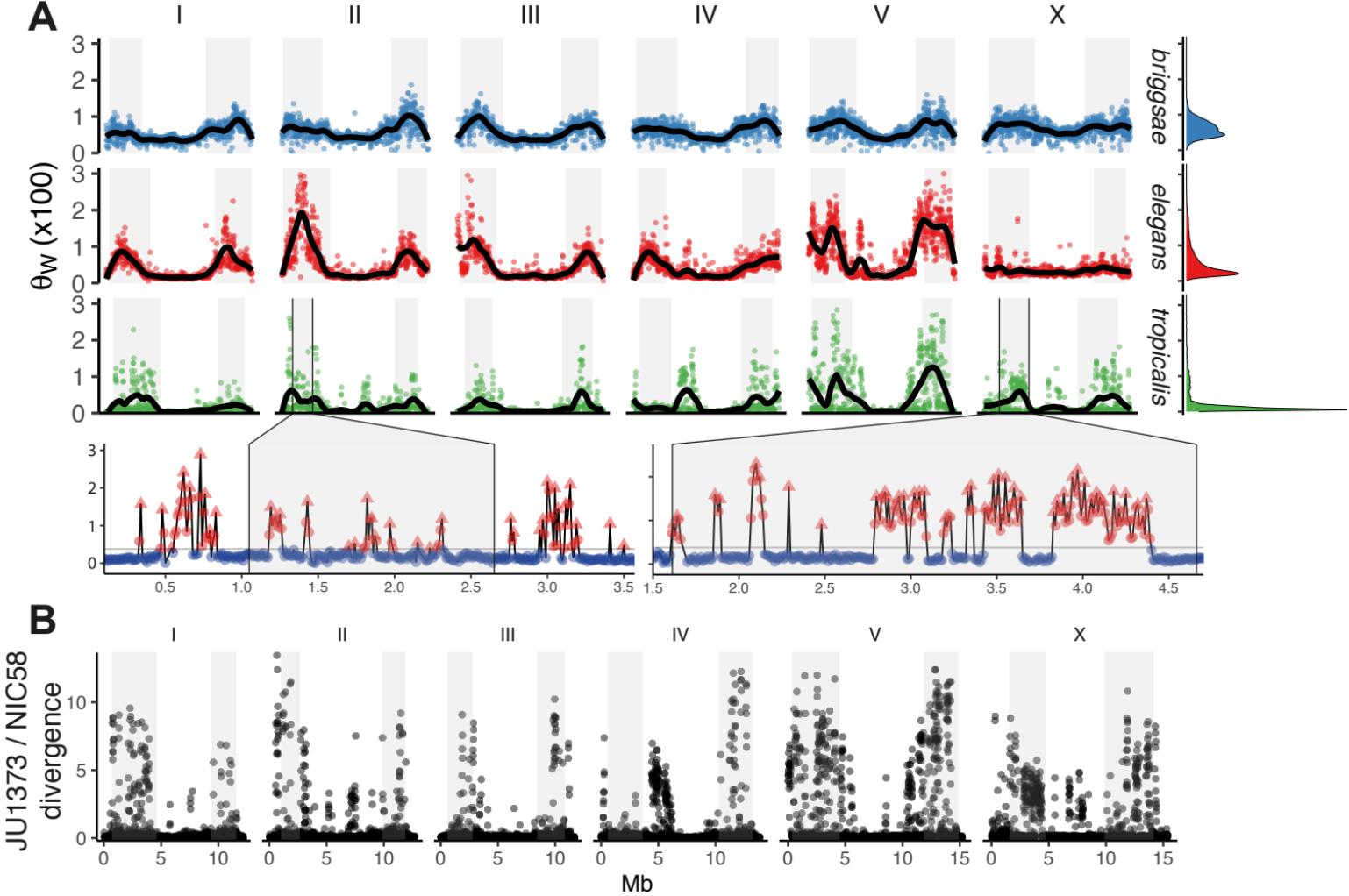
- 1665 21. <https://doi.org/10.1186/s12915-014-0093-1>
- 1666 Thomas, C. G., Wang, W., Jovelin, R., Ghosh, R., Lomasko, T., Trinh, Q., Kruglyak, L., Stein, L. D.,  
1667 & Cutter, A. D. (2015). Full-genome evolutionary histories of selfing, splitting, and selection in  
1668 *Caenorhabditis*. *Genome Research*, 25(5), 667–678. <https://doi.org/10.1101/gr.187237.114>
- 1669 Thompson, O. A., Snoek, L. B., Nijveen, H., Sterken, M. G., Volkers, R. J. M., Brenchley, R., Van't  
1670 Hof, A., Bevers, R. P. J., Cossins, A. R., Yanai, I., Hajnal, A., Schmid, T., Perkins, J. D.,  
1671 Spencer, D., Kruglyak, L., Andersen, E. C., Moerman, D. G., Hillier, L. W., Kammenga, J. E., &  
1672 Waterston, R. H. (2015). Remarkably Divergent Regions Punctuate the Genome Assembly of  
1673 the *Caenorhabditis elegans* Hawaiian Strain CB4856. *Genetics*, 200(3), 975–989.
- 1674 Ting, J. J., Woodruff, G. C., Leung, G., Shin, N.-R., Cutter, A. D., & Haag, E. S. (2014). Intense  
1675 Sperm-Mediated Sexual Conflict Promotes Reproductive Isolation in *Caenorhabditis*  
1676 Nematodes. *PLoS Biology*, 12(7), e1001915. <https://doi.org/10.1371/journal.pbio.1001915>
- 1677 van den Brand, T. (2020). *ggh4x: Hacks for “ggplot2.”*
- 1678 van der Kooi, C. J., & Schwander, T. (2014). On the fate of sexual traits under asexuality. *Biological  
1679 Reviews of the Cambridge Philosophical Society*, 89(4), 805–819.
- 1680 <https://doi.org/10.1111;brv.12078>
- 1681 Vaser, R., & Šikić, M. (2019). *Yet another de novo genome assembler* (p. 656306).  
1682 <https://doi.org/10.1101/656306>
- 1683 Vasimuddin, Md, Misra, S., Li, H., & Aluru, S. (2019). Efficient Architecture-Aware Acceleration of  
1684 BWA-MEM for Multicore Systems. In *arXiv [cs.DC]*. arXiv. <http://arxiv.org/abs/1907.12931>
- 1685 Wade, M. J., & Beeman, R. W. (1994). The population dynamics of maternal-effect selfish genes.  
1686 *Genetics*, 138(4), 1309–1314. <https://www.ncbi.nlm.nih.gov/pubmed/7896109>
- 1687 Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng,  
1688 Q., Wortman, J., Young, S. K., & Earl, A. M. (2014). Pilon: an integrated tool for comprehensive  
1689 microbial variant detection and genome assembly improvement. *PloS One*, 9(11), e112963.  
1690 <https://doi.org/10.1371/journal.pone.0112963>
- 1691 Weichenhan, D., Kunze, B., Traut, W., & Winking, H. (1998). Restoration of the Mendelian

- 1692 transmission ratio by a deletion in the mouse chromosome 1 HSR. *Genetical Research*, 71(2),  
1693 119–125. <https://doi.org/10.1017/s0016672398003206>
- 1694 Weichenhan, D., Traut, W., Kunze, B., & Winking, H. (1996). Distortion of Mendelian recovery ratio  
1695 for a mouse HSR is caused by maternal and zygotic effects. *Genetical Research*, 68(2), 125–  
1696 129. <https://doi.org/10.1017/s0016672300034017>
- 1697 Wei, Q., Zhao, Y., Guo, Y., Stomel, J., Stires, R., & Ellis, R. E. (2014). Co-option of alternate sperm  
1698 activation programs in the evolution of self-fertile nematodes. *Nature Communications*, 5, 5888.  
1699 <https://doi.org/10.1038/ncomms6888>
- 1700 Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.  
1701 <https://ggplot2.tidyverse.org>
- 1702 Wickham, H., François, R., Henry, L., & Müller, K. (2020). *dplyr: A Grammar of Data Manipulation*.  
1703 <https://CRAN.R-project.org/package=dplyr>
- 1704 Wickham, H., & Henry, L. (2020). *tidy: Tidy Messy Data*. <https://CRAN.R-project.org/package=tidy>
- 1705 Wick, R. R., Judd, L. M., Gorrie, C. L., & Holt, K. E. (2017). Unicycler: Resolving bacterial genome  
1706 assemblies from short and long sequencing reads. *PLoS Computational Biology*, 13(6),  
1707 e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
- 1708 Wilson, D. S., & Colwell, R. K. (1981). Evolution of Sex Ratio in Structured Demes. *Evolution;*  
1709 *International Journal of Organic Evolution*, 35(5), 882–897. <https://doi.org/10.2307/2407858>
- 1710 Winking, H., Weith, A., Boldyreff, B., Moriwaki, K., Fredga, K., & Traut, W. (1991). Polymorphic  
1711 HSRs in chromosome 1 of the two semispecies *Mus musculus musculus* and *M. m. domesticus*  
1712 have a common origin in an ancestral population. *Chromosoma*, 100(3), 147–151.  
1713 <https://doi.org/10.1007/BF00337242>
- 1714 Yin, D., Schwarz, E. M., Thomas, C. G., Felde, R. L., Korf, I. F., Cutter, A. D., Schartner, C. M.,  
1715 Ralston, E. J., Meyer, B. J., & Haag, E. S. (2018). Rapid genome shrinkage in a self-fertile  
1716 nematode reveals sperm competition proteins. *Science*, 359(6371), 55–61.  
1717 <https://doi.org/10.1126/science.aoa0827>
- 1718 Zhao, Y., Tan, C.-H., Krauchunas, A., Scharf, A., Dietrich, N., Warnhoff, K., Yuan, Z., Druzhinina,

1719 M., Gu, S. G., Miao, L., Singson, A., Ellis, R. E., & Kornfeld, K. (2018). The zinc transporter  
1720 ZIPT-7.1 regulates sperm activation in nematodes. *PLoS Biology*, 16(6), e2005069.  
1721 <https://doi.org/10.1371/journal.pbio.2005069>

1722





**A**

CV : Cape Verde (1)

R : La Réunion (1)

STP : São Tomé and Príncipe (1)

B : Brazil (1)

CR : Costa Rica (1)

FG : French Guiana (3)

Pa : Panama (1)

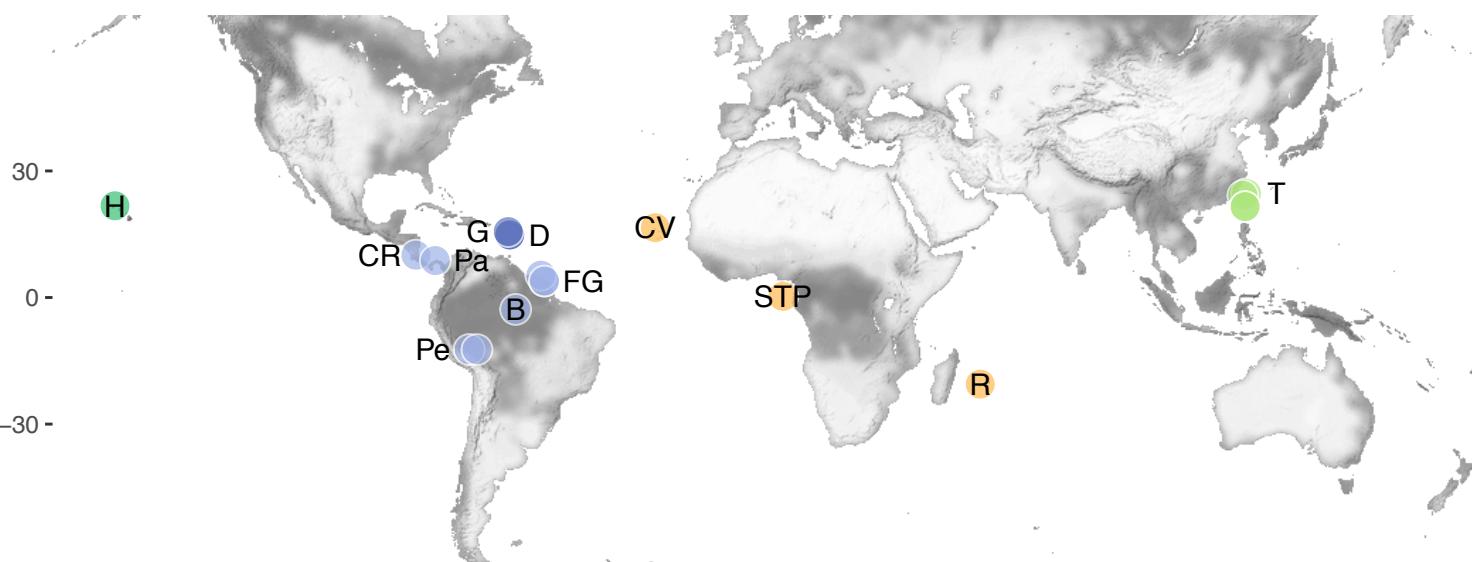
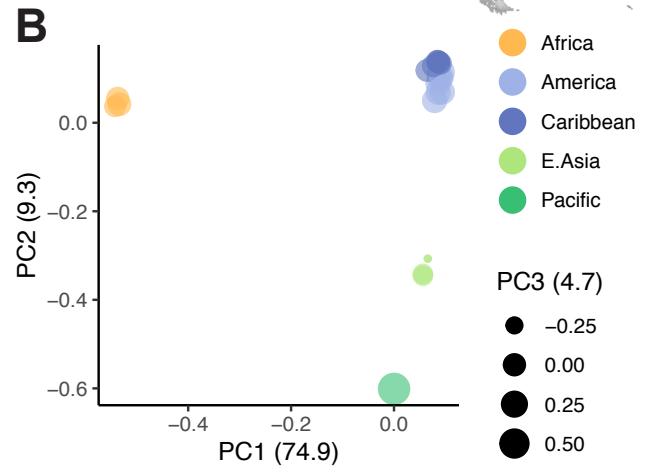
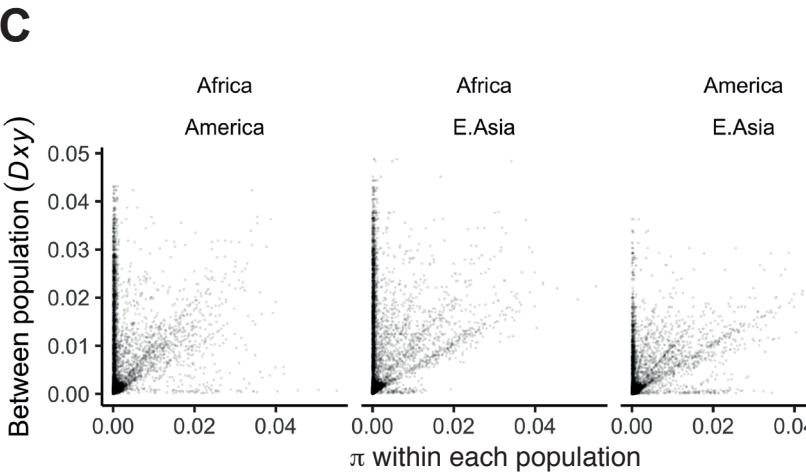
Pe : Peru (2)

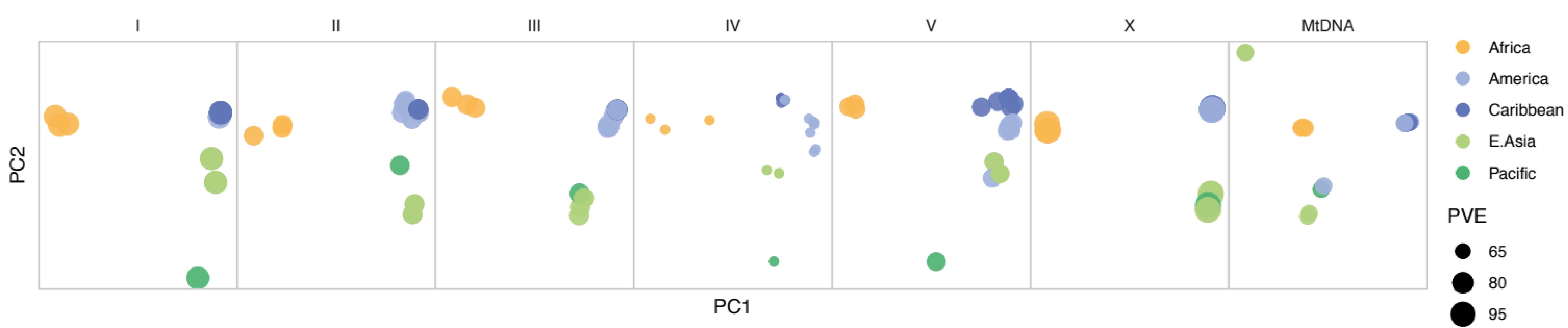
D : Dominica (7)

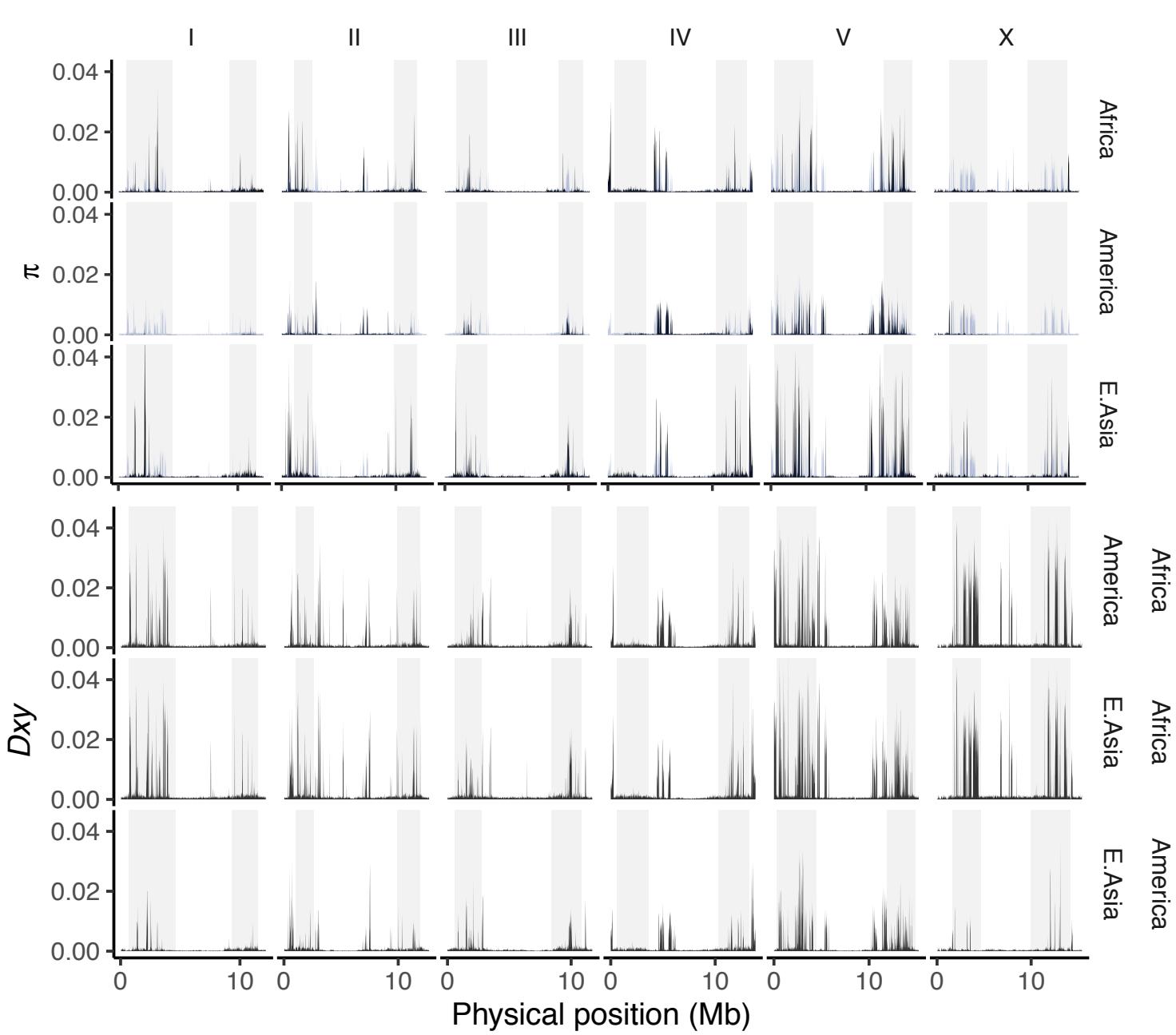
G : Guadeloupe (1)

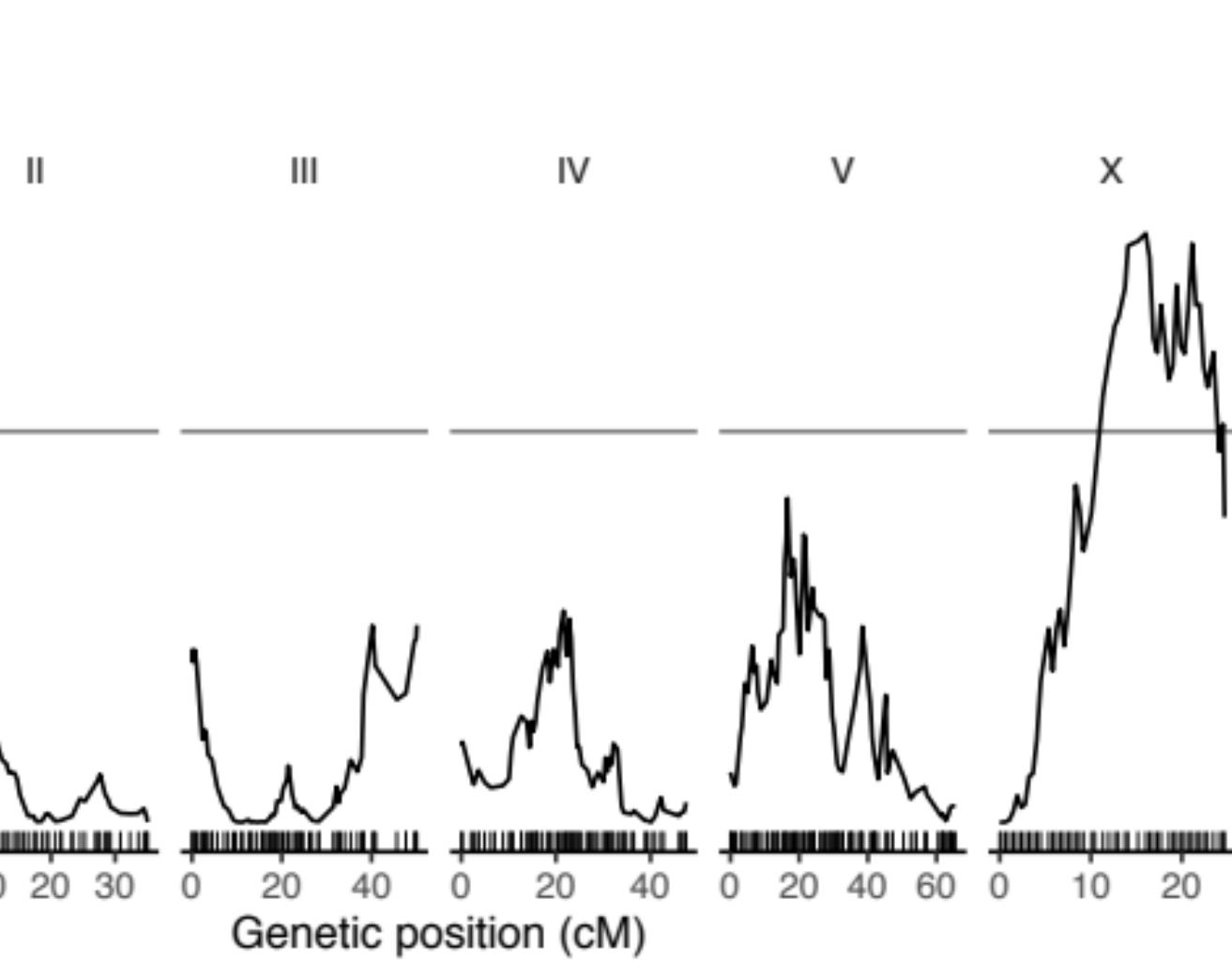
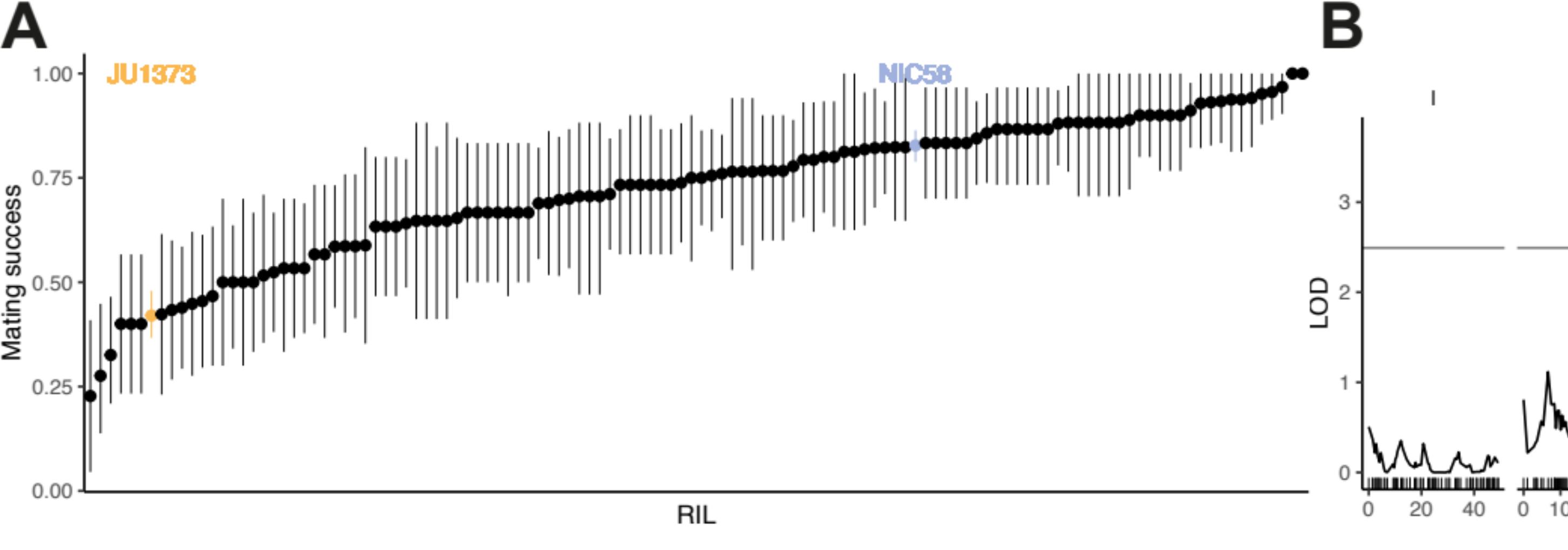
T : Taiwan (4)

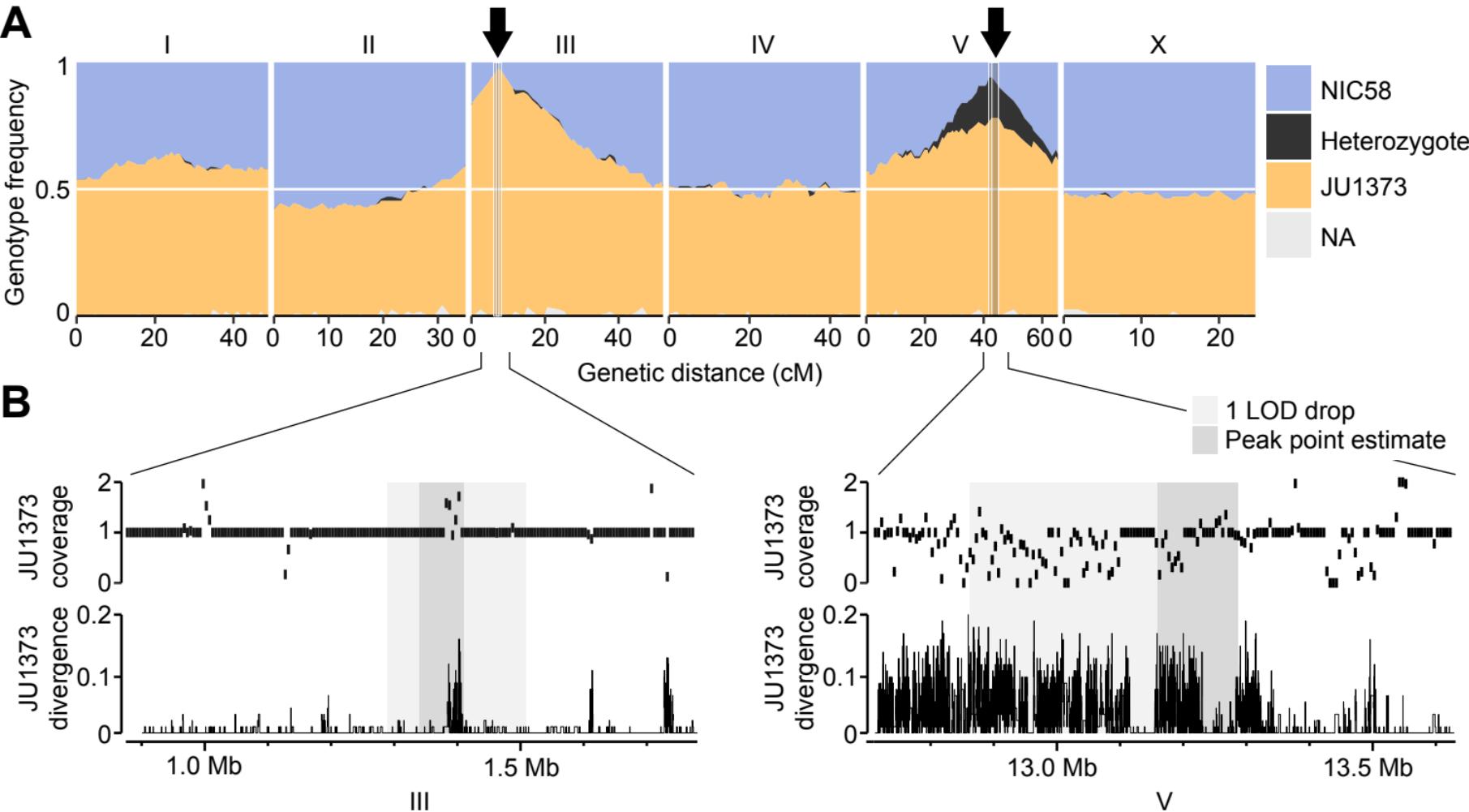
H : Hawaii (1)

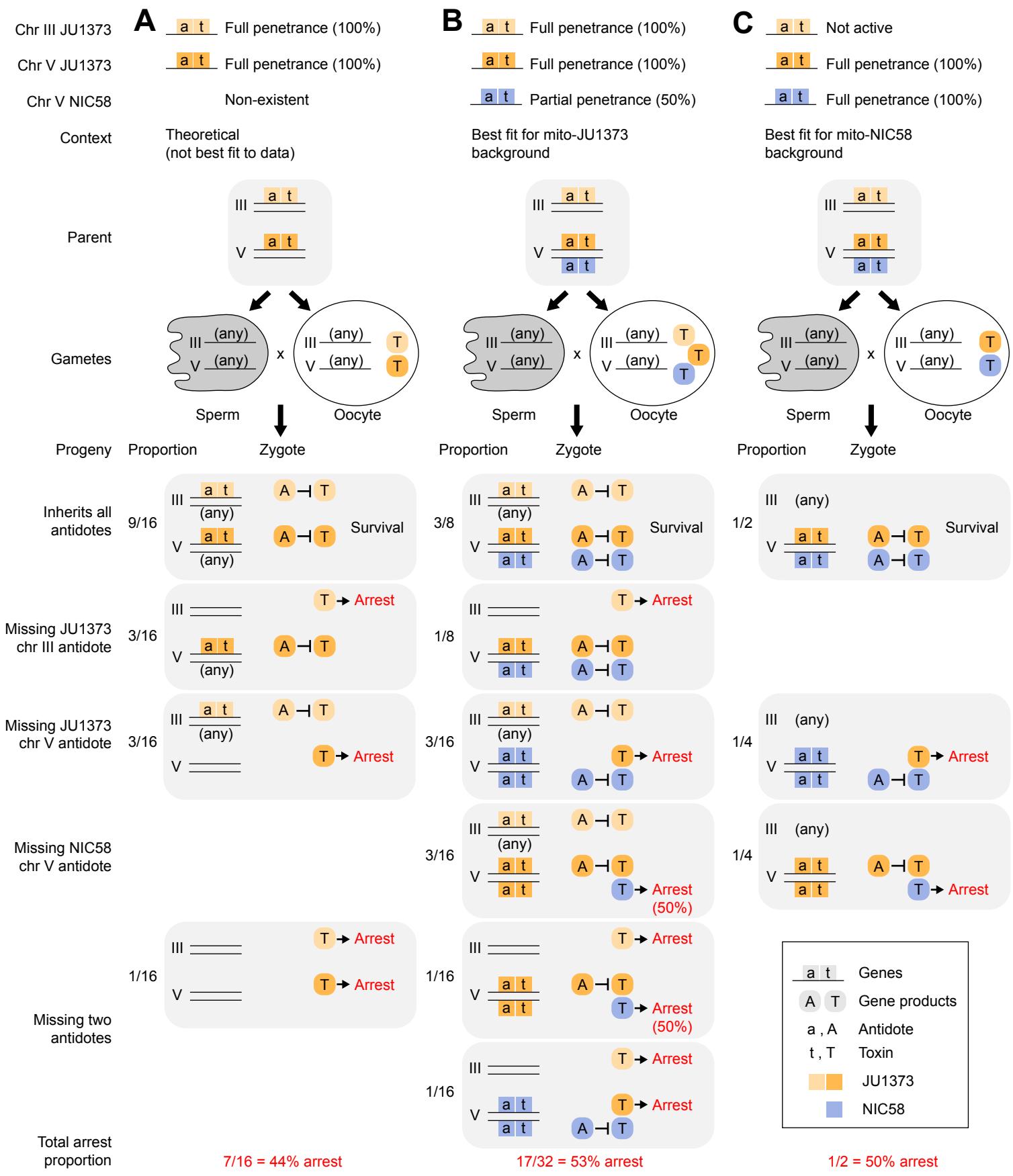
**B****C**

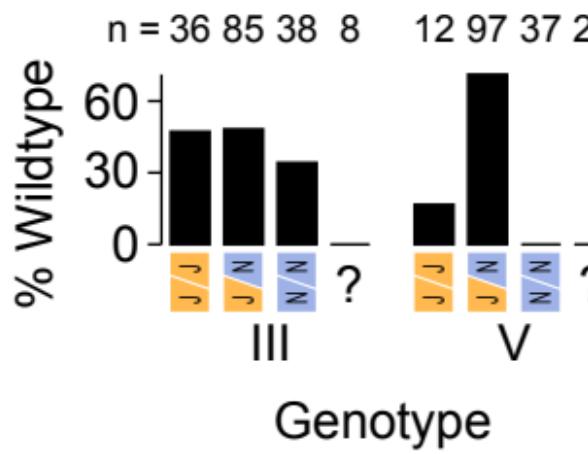
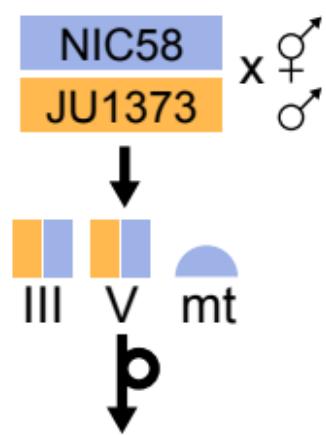
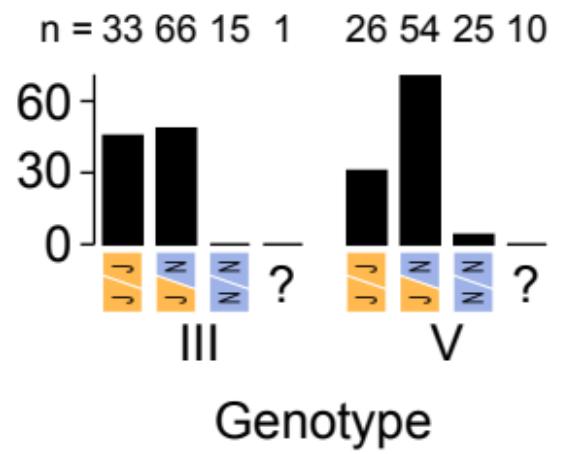
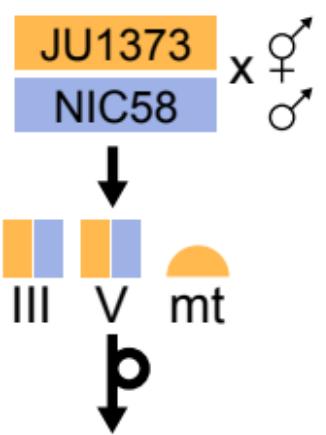
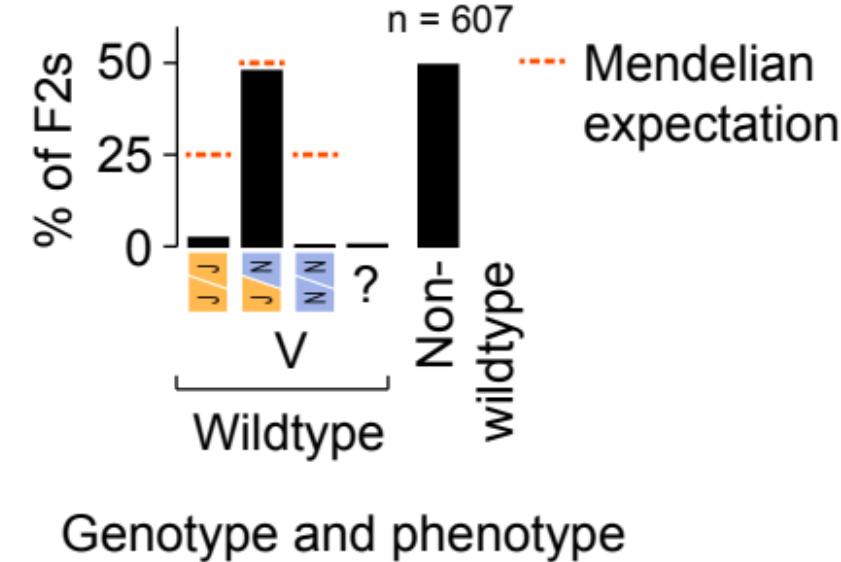
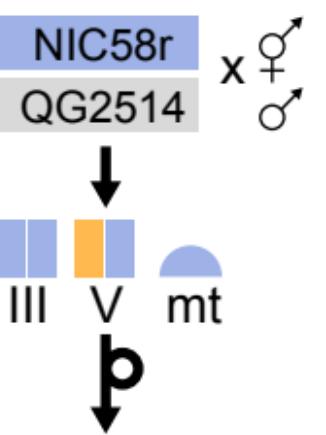
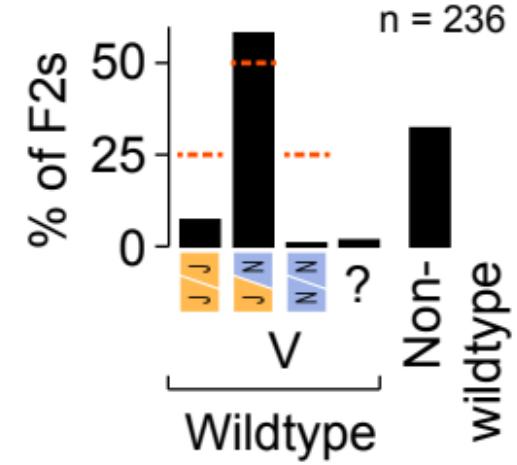
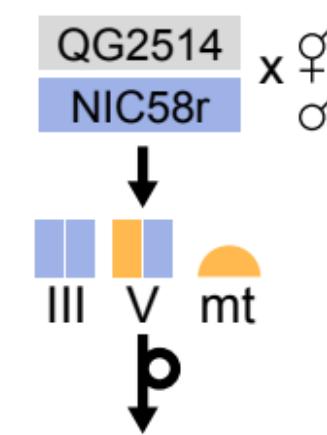


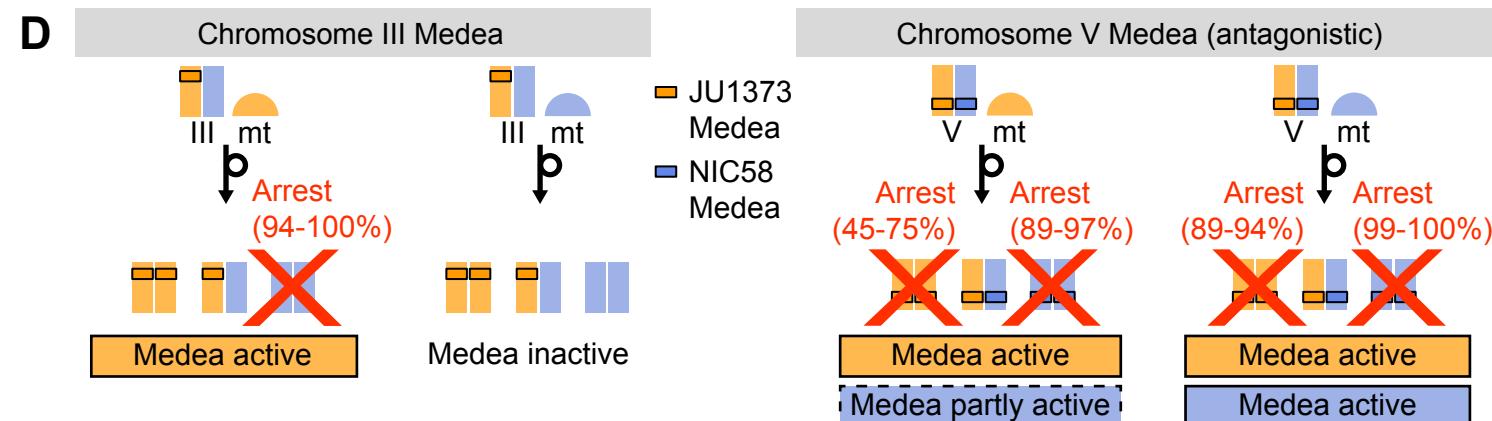
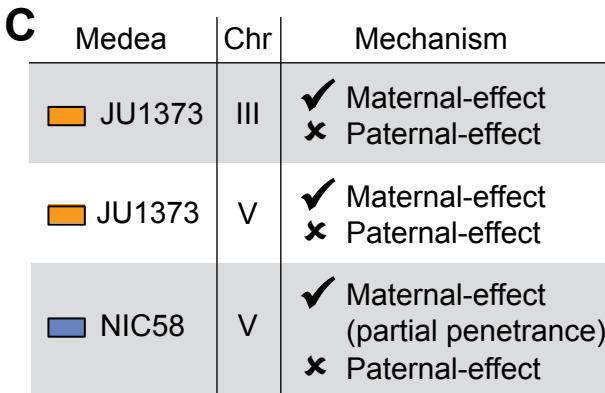
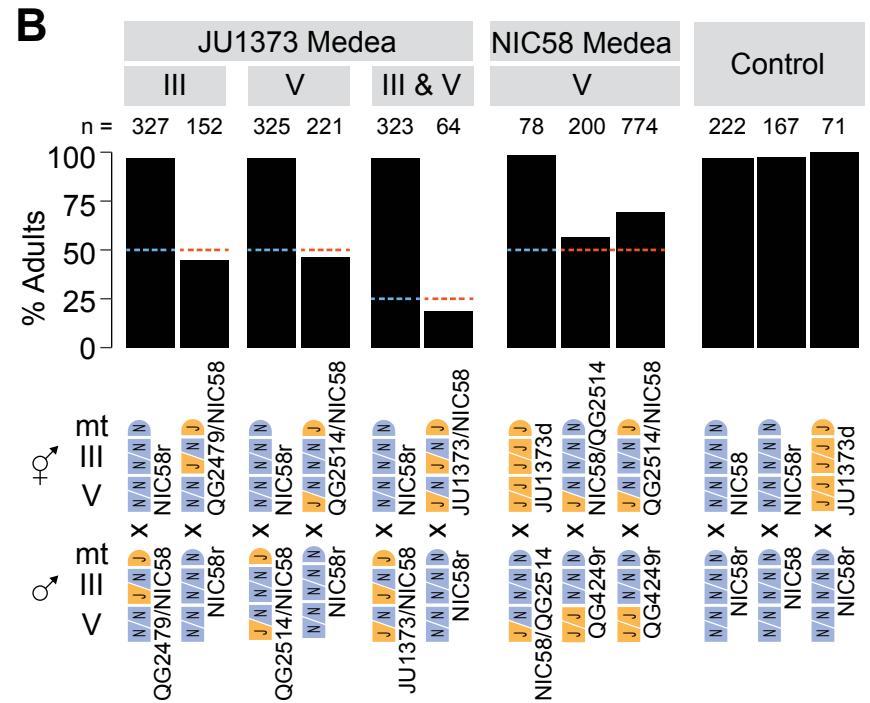
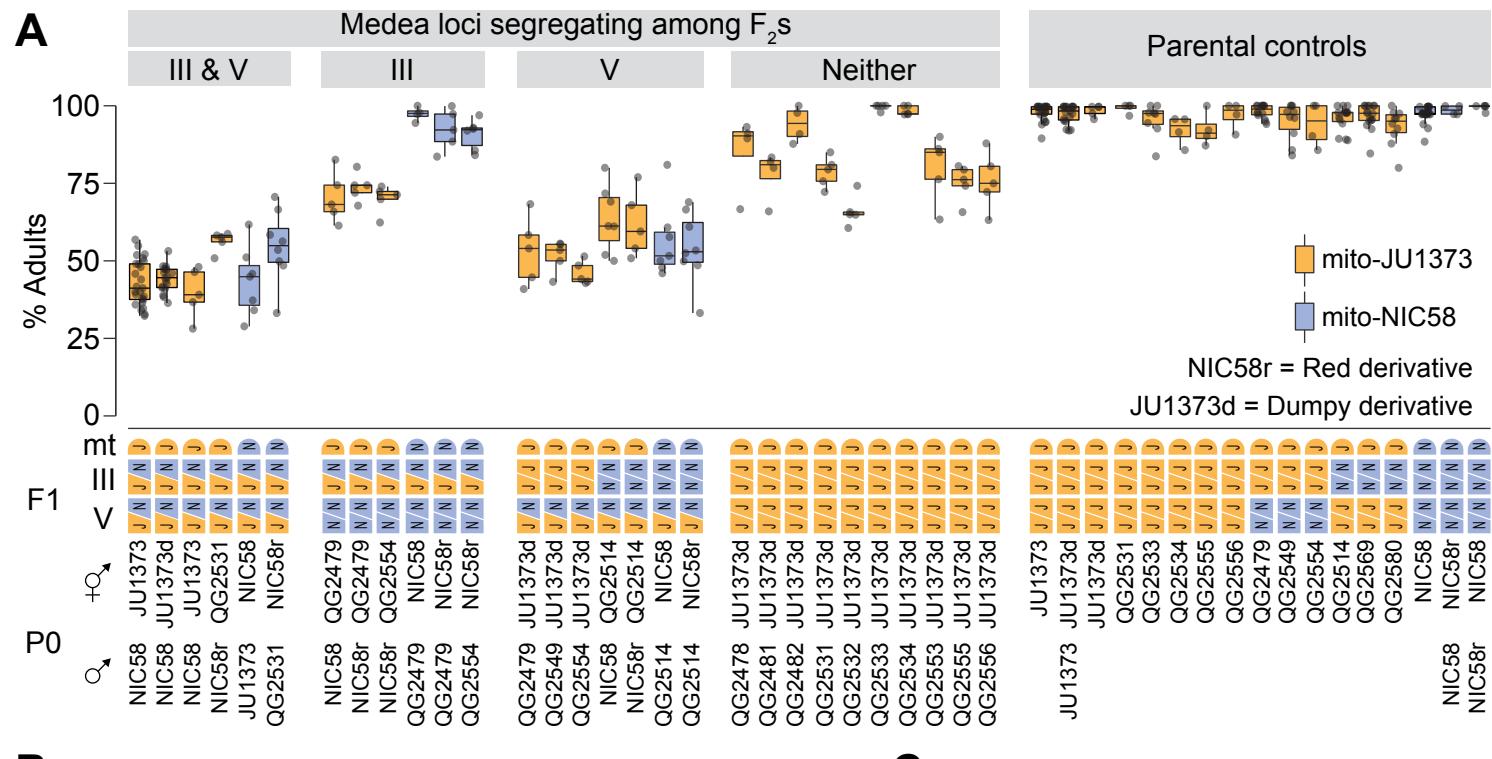


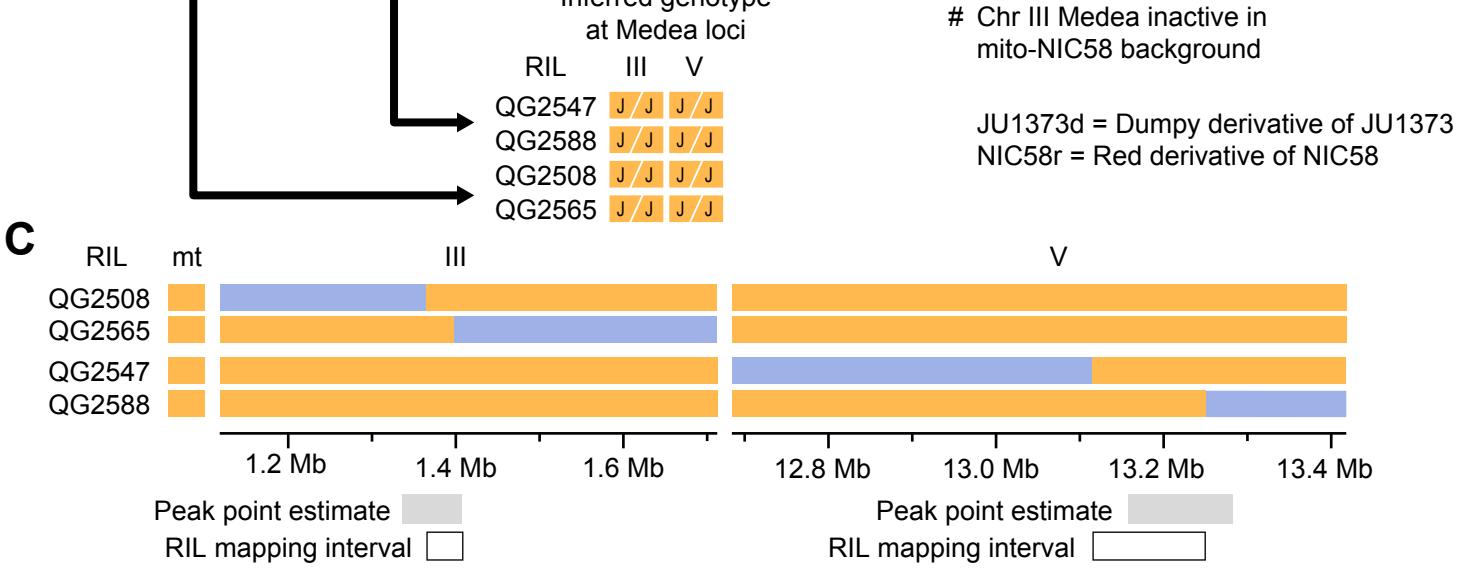
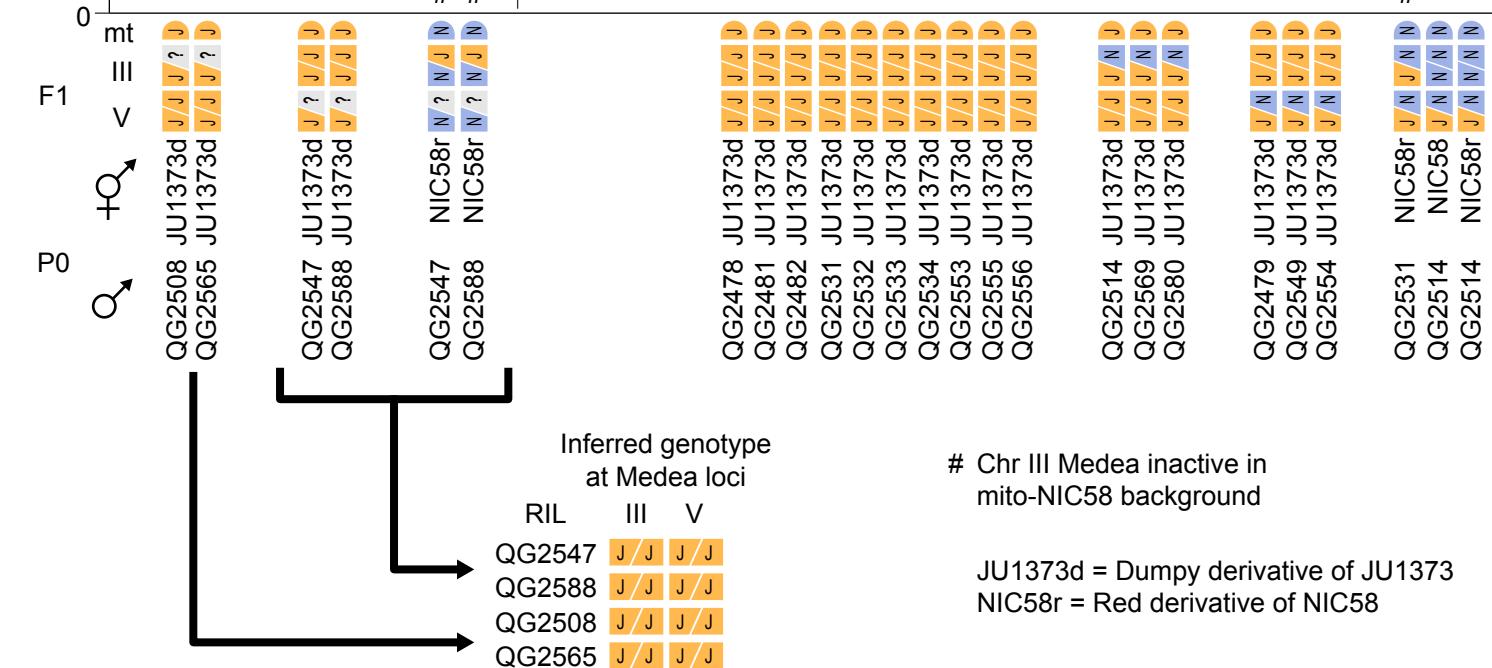
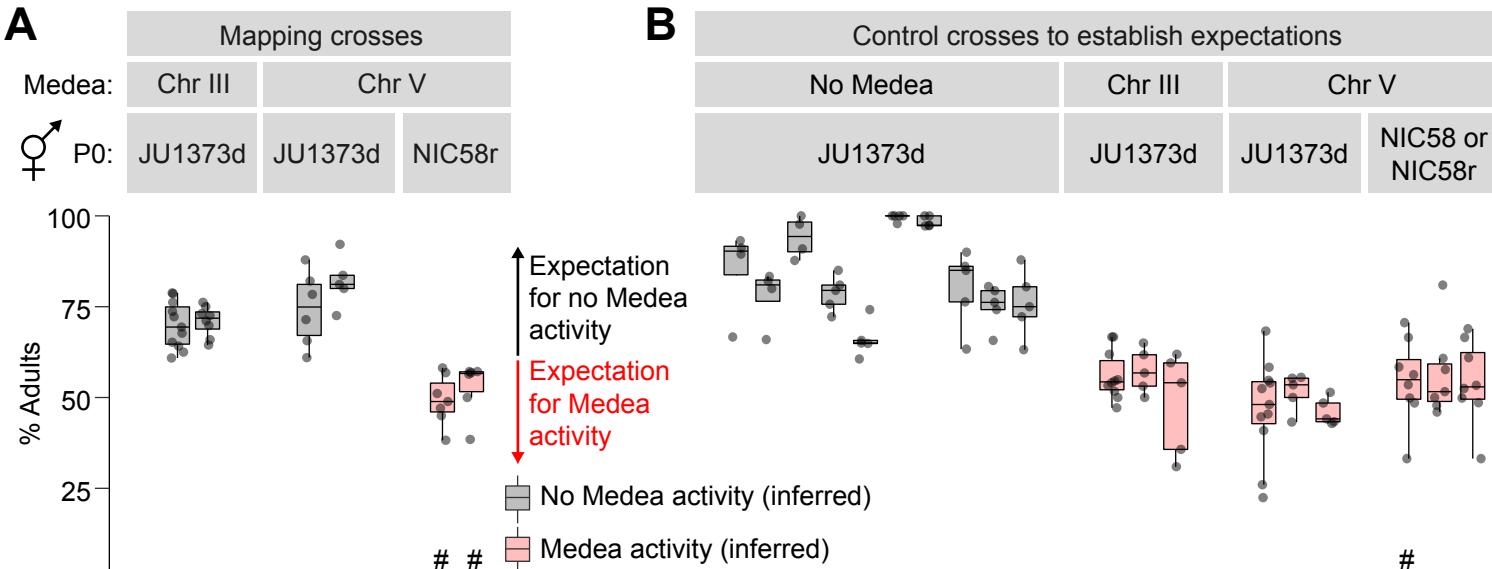


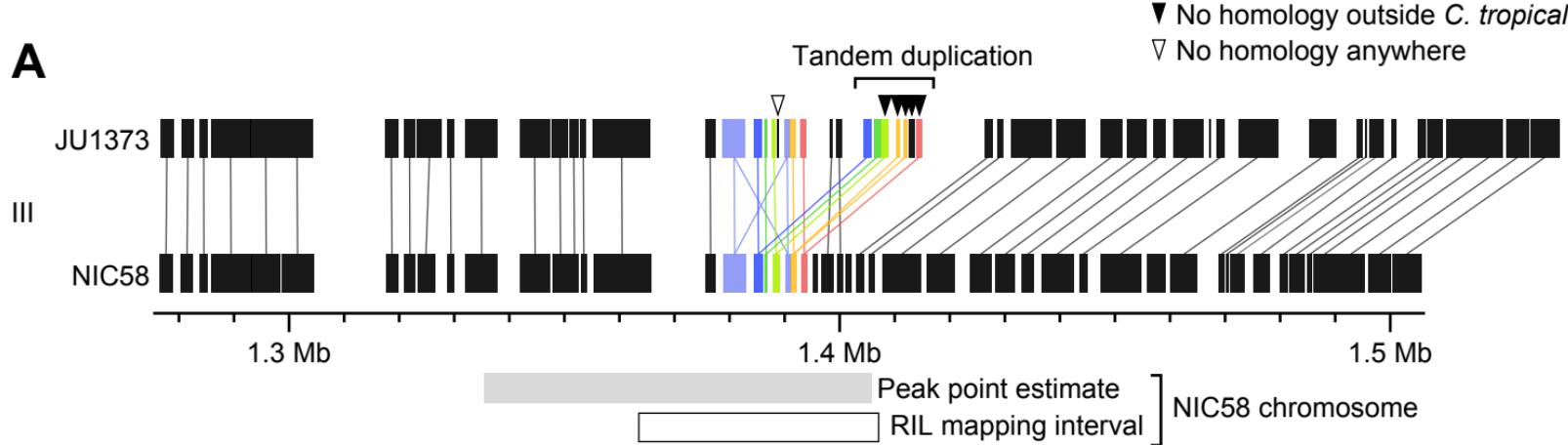
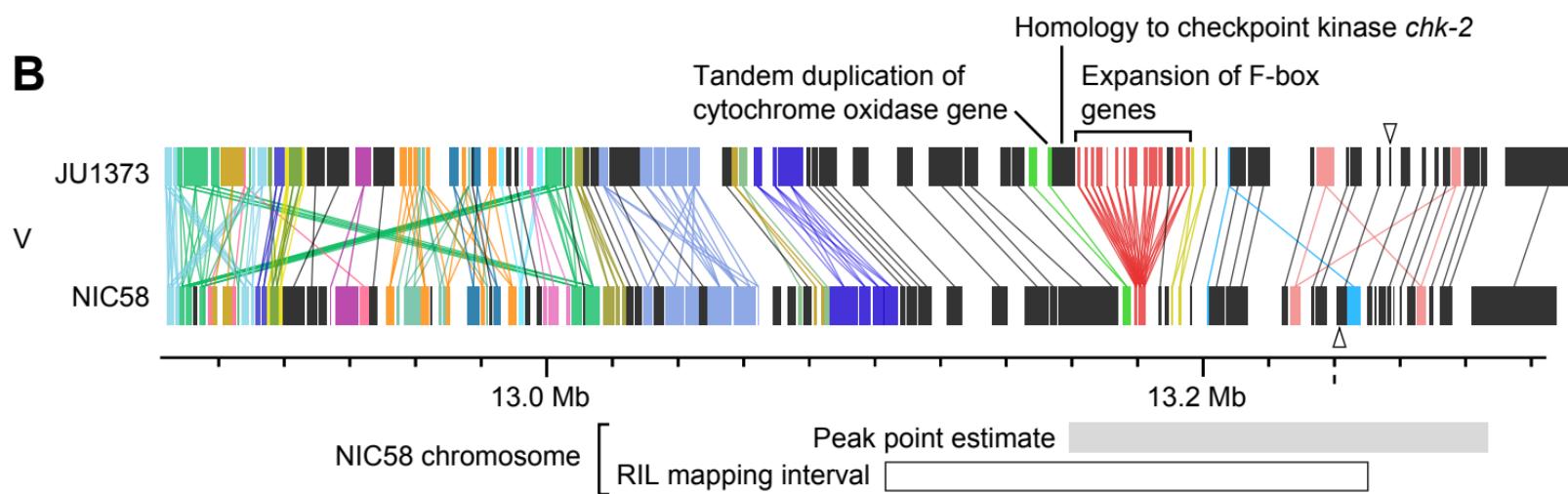


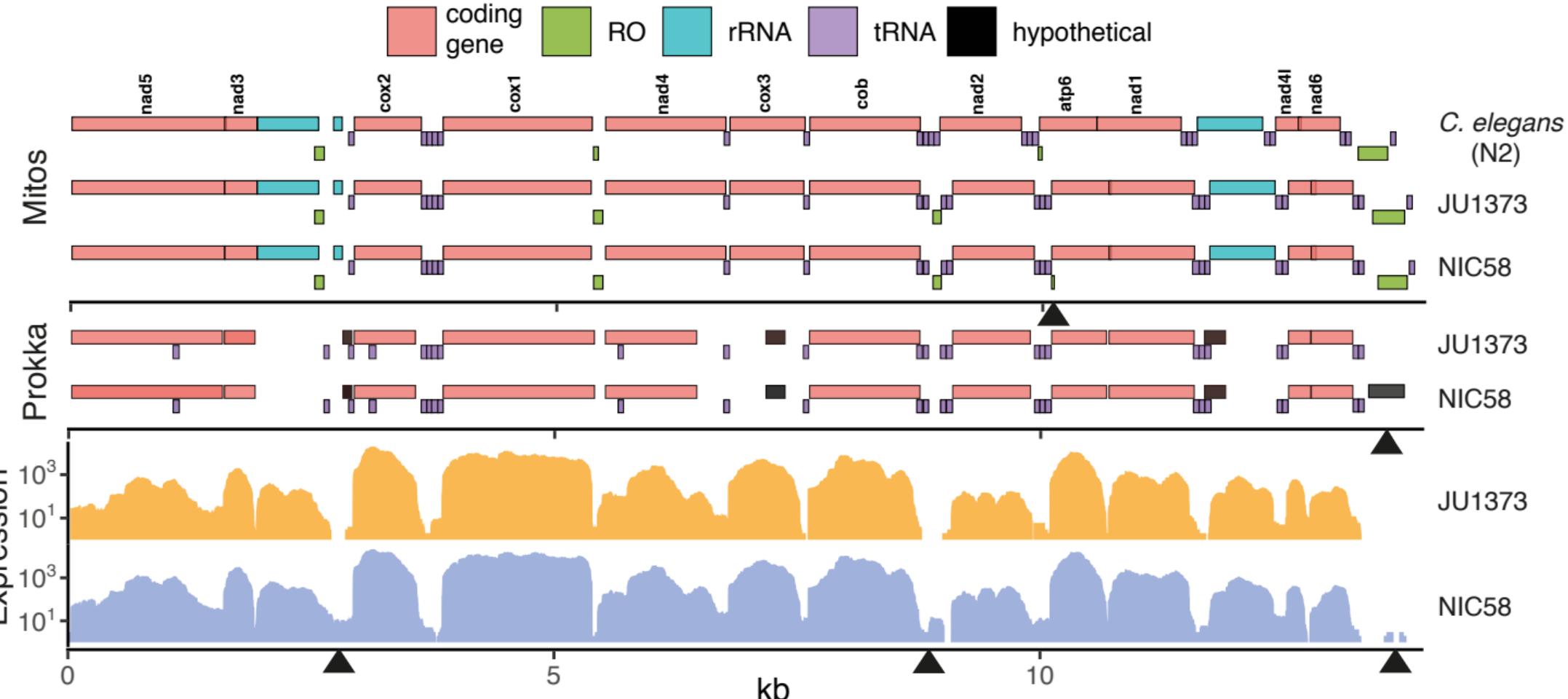


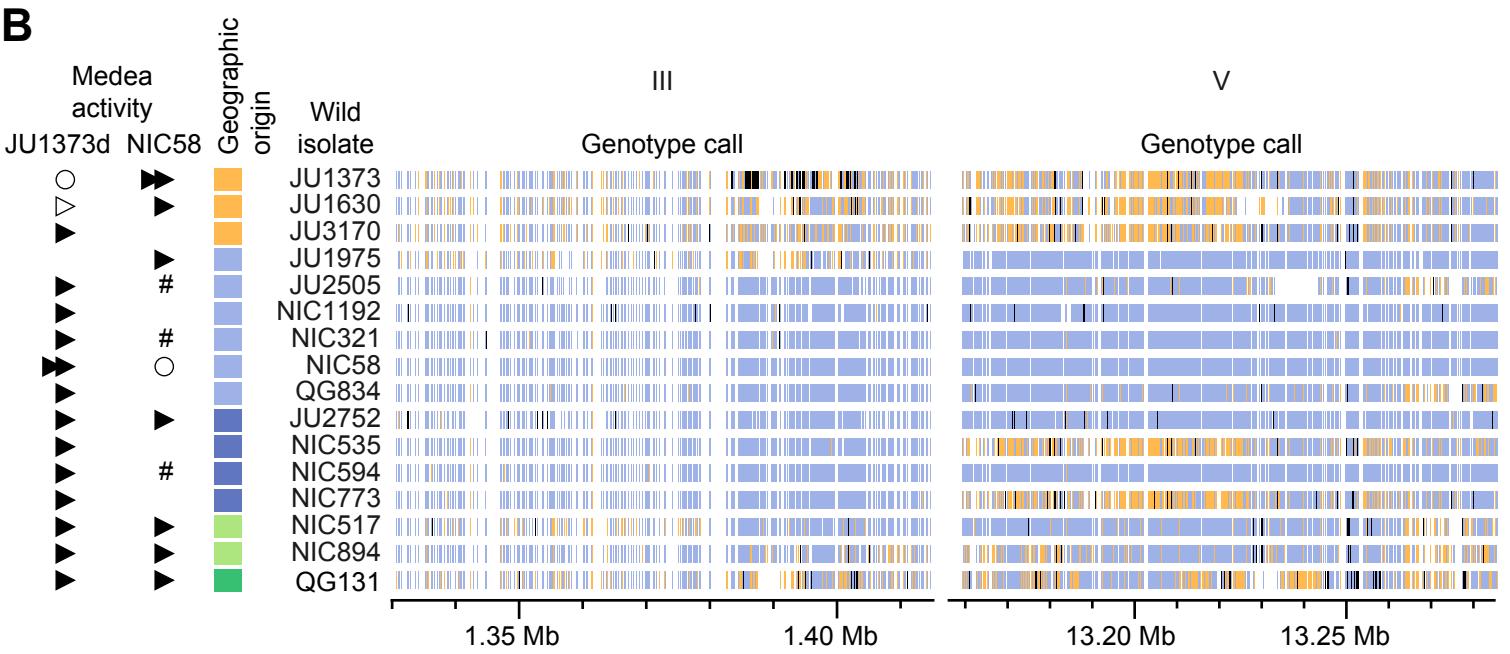
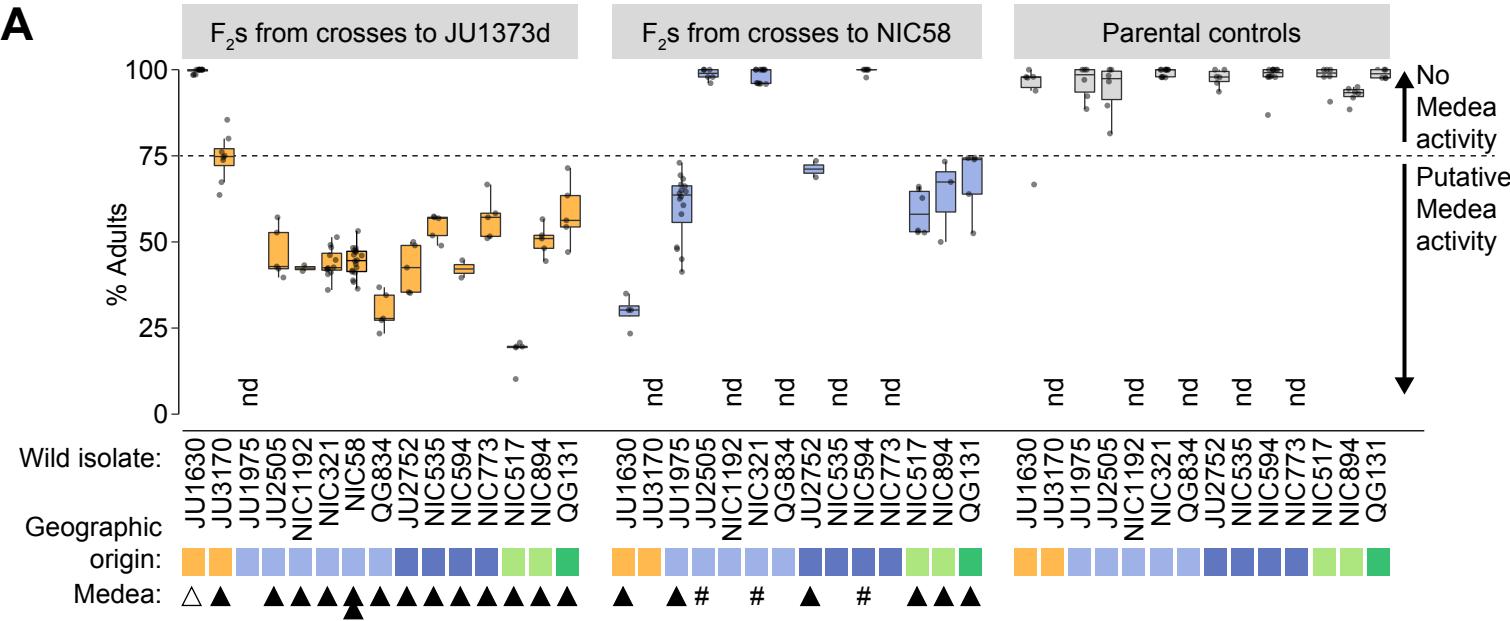
**A****B**



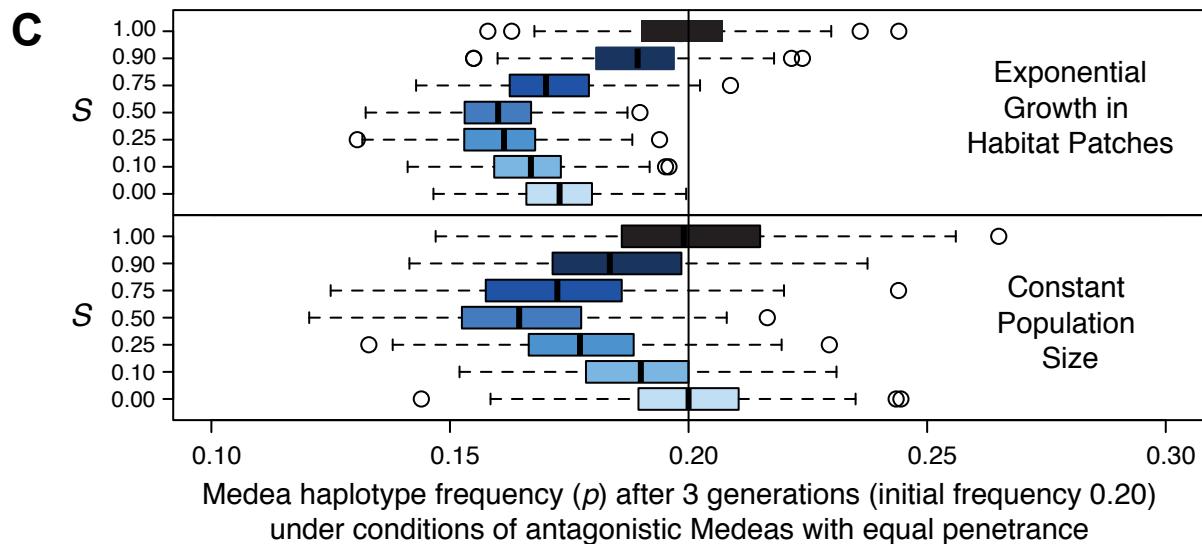
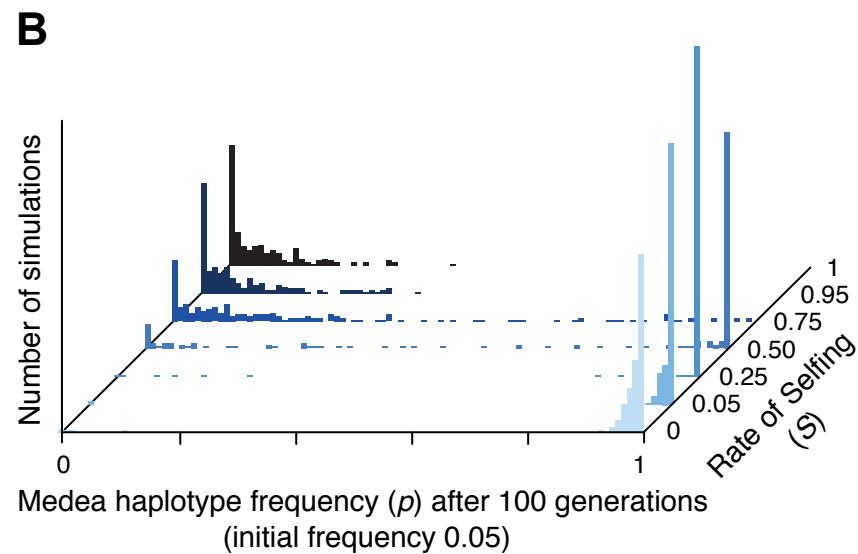
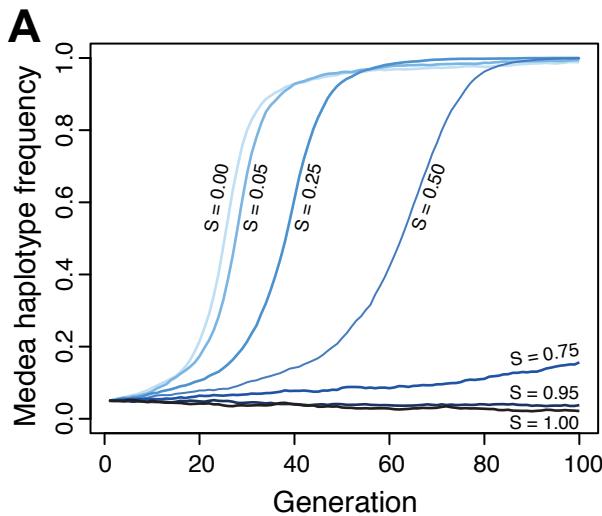


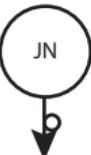
**A****B**





Medea activity	Geographic origin	Genotype call
► NIC58 x JU1373 Medea	► Africa	► Reference allele (NIC58)
► Putative Medea activity	► Americas	► Heterozygous
▷ No Medea activity	► Caribbean	► Alternate allele
○ Self	► East Asia	
# Inconclusive	► Pacific	





50% J

JJ	JJ
JJ	JJ

100% J

JJ	JN
JN	

2/3 J

NN	NN
NN	NN

0% J

### Genic Selection

Medea increases J frequency in the progeny of heterozygotes

54% J

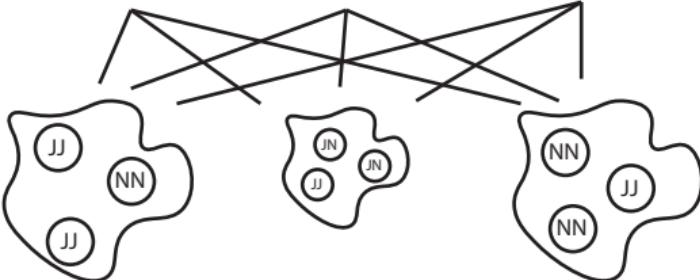
100 progeny



75 progeny



100 progeny



6 days to  
dauer

10 days to  
dauer

6 days to  
dauer

### Group Selection

Populations in habitat patches colonized by fewer heterozygotes will grow more rapidly and disperse to new patches sooner

