Caenorhabditis hybridizations reveal cytoplasmic-nuclear incompatibilities and asexual reproduction

Piero Lamelza

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Michael Ailion, Chair

Harmit S. Malik

Maitreya J. Dunham

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

Abstract

Caenorhabditis hybridizations reveal cytoplasmic-nuclear incompatibilities and asexual reproduction

Piero Lamelza

Chair of the Supervisory Committee:
Michael Ailion
Department of Biochemistry

Part 1:

How species arise is a fundamental question in biology. Species can be defined as populations of interbreeding individuals that are reproductively isolated from other such populations. Therefore, understanding how reproductive barriers evolve between populations is essential for understanding the process of speciation. Hybrid incompatibility (for example, hybrid sterility or lethality) is a common and strong reproductive barrier in nature. Here we report a lethal incompatibility between two wild isolates, NIC59 and JU1825, of the nematode *Caenorhabditis nouraguensis*. Hybrid inviability results from the incompatibility between a maternally inherited cytoplasmic factor from each strain and a recessive nuclear locus from the other. Furthermore, cytoplasmic-nuclear incompatibility commonly occurs between other wild isolates, indicating that this is a significant reproductive barrier within *C. nouraguensis*. We hypothesize that the maternally inherited cytoplasmic factor is the mitochondrial genome and that mitochondrial dysfunction underlies

hybrid death. We find that the JU1825 nuclear incompatibility locus maps to chromosome III and the NIC59 nuclear incompatibility locus maps to chromosome IV, indicating the incompatibilities are genetically distinct. We have finely mapped the JU1825 nuclear incompatibility locus to a 100 kb region containing tandem and diversified repeats of F-box domain, transmembrane domain and nuclear hormone receptors genes. There are also exact tandem repeats of a leucine-tRNA. The NIC59 and JU1825 mitochondrial genomes differ by 95 SNPs and 1 bp indel. Potential functional differences between the genomes include single non-synonymous changes in *nd-1* and *cox-1* (components of complex I and IV of the electron transport chain, respectively), a SNP in the cysteine tRNA and 3 SNPs and 1 bp indel in the 16s rRNA. This system has the potential to shed light on the dynamics of divergent mitochondrial-nuclear coevolution and its role in promoting speciation.

Part 2:

Most animal species reproduce by sex. Theory predicts there are advantages to being able to switch reproduction between sexual and asexual modes. However, facultative sex is rarely observed in animals, implying that there are strong selective pressures that prevent asexuality arising from an obligately sexual ancestor. One of the critical steps in the evolution of asexuality from a sexual ancestor is the transition from haploid to diploid maternal inheritance. Here we report that interspecific hybridization between two sexual *Caenorhabditis* nematode species (*C. nouraguensis* females and *C. becei* males) results in two classes of viable offspring. The first class consists of fertile offspring, which are produced asexually by sperm-dependent parthenogenesis (also called gynogenesis or pseudogamy); these progeny inherit a diploid maternal genome but fail to inherit a paternal genome. The second class consists of sterile hybrid

offspring, which inherit both a diploid maternal genome and a haploid paternal genome. Using whole-genome sequencing of individual viable worms, we show that diploid maternal inheritance in both asexually produced and hybrid offspring results from the inheritance of two randomly selected homologous chromatids from *C. nouraguensis* oocytes. This genetic mechanism of diploid maternal inheritance is indistinguishable from that of many obligately asexual species. Furthermore, we show that intraspecies *C. nouraguensis* crosses can also result in a low frequency of asexual reproduction through diploid maternal inheritance. Thus, *C. nouraguensis* provides a genetically tractable model to study the evolutionary origins of asexuality from obligately sexual species.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I thank my mother, Eduviges Lamelza, and my father, Vincent M. Lamelza, for raising me and loving me unconditionally. I thank my brother, Paolo Lamelza, for engaging conversations and arguments and my aunt, Guillermina Gomez, for her humor and friendship.

I would like to thank my thesis advisor, Michael Ailion, for his mentorship and never-ending interest in weird biology (e.g. hermaphroditic beavers). I would also like to thank Irini Topalidou for her mentorship and her kind words that have slowly have built-up my confidence as a scientist. Most importantly, Michael and Irini have shown me that curiosity, creativity and good science are not limited to "famous" labs.

Most of my thesis work was done in close collaboration with Janet M. Young and Harmit S. Malik of the Fred Hutchinson Cancer Research Center. I am grateful for their eagerness to collaborate and their expertise in computational and evolutionary biology, which has elevated this work beyond what I could have done on my own. I also would like to acknowledge Mark Blaxter, Lewis Stevens and everyone working on the Caenorhabditis Genomes Project for generating the DNA and RNA sequencing data used to generate a reference genome assembly for *C. nouraguensis*. I also thank Luke Noble and Matthew V. Rockman for graciously sharing their *C. becei* genome assembly. I am also grateful for the advice and attention of my committee members: Harmit S. Malik, Maitreya J. Dunham, James R. Priess and Jay Shendure.

DEDICATION

I dedicate this work to my girlfriend Emily A. Scarborough (a.k.a. huevo malo, gato malo and BB). Drinking beer and talking smack with her has been the best part of graduate school. I am a better person because of her.

Chapter 1. INTRODUCTION

While working in my thesis lab, I was primarily interested in understanding the genetic basis of hybrid incompatibility and its relation to the process of speciation. While studying hybridizations between different wild-isolates of *Caenorhabditis nouraguensis* nematodes, I discovered a wide-spread genetic incompatibility that results in hybrid death. I then attempted to identify the genetic components of the incompatibility. The first section of my introduction, entitled "The genetic basis of hybrid incompatibility", gives context for the goals of this project.

While studying interspecies hybridizations between two *Caenorhabditis* sister-species, *C. nouraguensis* and *C. becei*, I discovered that *C. nouraguensis* females can reproduce asexually at a low frequency. This led to a project that focuses on understanding the evolutionary transition between sexual and asexual reproduction. The second section of my introduction, entitled "The evolution of asexuality from a sexual ancestor", gives context for the goals of this project.

1.1 THE GENETIC BASIS OF HYBRID INCOMPATIBILITY

How species arise is a fundamental and still unanswered question in biology. Darwin's *On the Origin of Species* was the first major work to recognize that new species emerge from those already in existence. Specifically, Darwin viewed species as morphologically discrete groups that evolved from a common ancestor by differential ecological adaptation. However, this view was later seen by evolutionary biologists as incomplete. The rediscovery of Mendel's work and the birth of population genetics made it apparent that genetically based morphological differences between populations could not be maintained in the face of gene flow. Therefore, reproductive barriers that reduce gene flow were considered to be essential in promoting divergence and allowing new species to remain distinct and undergo unique evolutionary

histories. Accordantly, Ernst Mayr promoted a biological species concept based on the ability of populations to interbreed and exchange genetic information rather than their morphological similarities. Specifically, Mayr defined species as populations of actually or potentially interbreeding individuals that are reproductively isolated from other such populations (Mayr 1942). Under this definition, discovering how new species are formed requires understanding how reproductive barriers evolve between populations.

Reproductive barriers are often binned into either one of two categories: pre-zygotic or post-zygotic. Pre-zygotic barriers prevent the fusion of heterospecific gametes and include differences in behavior, habitat, pollinators, breeding time and gamete compatibility. Post-zygotic barriers refer to the formation of hybrid zygotes that are relatively unfit in comparison to their parents and serve as inefficient bridges for gene flow between populations. Hybrids can be extrinsically unfit, in that they are maladapted to their environment (e.g. hybrids exhibit an intermediate phenotype which is unfit in parental environments) or intrinsically unfit, in that they are developmentally abnormal (for example, hybrids are sterile or inviable) (Coyne and Orr 2004). Species are typically separated by several reproductive barriers, making it difficult to distinguish which were instrumental in initially reducing gene flow and which evolved after speciation was complete. Therefore, the relative importance of any of the reproductive barriers mentioned above in driving speciation is largely unknown and most likely varies across taxa. However, a significant amount of research has focused on intrinsic post-zygotic barriers due to their relative ease of study in the laboratory in comparison to barriers that require an environmental context.

Theodosius Dobzhansky and Hermann Muller, two geneticists studying hybrid sterility and lethality in *Drosophila* in the early 20th century, hypothesized that hybrids were

developmentally unfit due to incompatible gene combinations (Dobzhansky 1933; Maheshwari and Barbash 2011). In its simplest form, their model (Dobzhansky-Muller incompatibility model) predicts that at least two genetic loci, each having evolved independently in one of two divergent lineages, have deleterious epistatic interactions in hybrids. This model has gained support by the molecular identification of genes required for hybrid dysfunction in several genera (Presgraves 2010). Identifying these genes and the natural forces that drive their divergence is one of the major objectives of speciation genetics.

As previously stated, Darwin suggested that differential ecological adaptation by natural selection was the major driving force for speciation. Some of the molecularly identified incompatibility genes do indeed show signs of selection (Ting *et al.* 1998; Presgraves *et al.* 2003; Barbash *et al.* 2004; Brideau *et al.* 2006; Oliver *et al.* 2009; Phadnis *et al.* 2015) with some likely driven by ecological adaptation (Chae *et al.* 2014). For example, a systematic screen for incompatibility loci underlying hybrid necrosis in *Arabidopsis thaliana* revealed that most map to genomic regions enriched for nucleotide-binding domain and leucine-rich repeat (NLR) immune receptor genes. Under normal conditions, NLRs bind to plant-pathogen effector proteins (either pathogen proteins or pathogen-modified host proteins) through their leucine-rich repeat (LRR) domain and then trigger the death of infected and surrounding cells (Steinbrenner *et al.* 2012). However, in hybrids NLRs activate cell death in the absence of pathogens. A majority of the NLRs present at the incompatibility loci are members of the *RPP1* subfamily, which recognize the oomycete pathogen *Hyaloperonospora arabidopsis* (*Hpa*). The *RPP* family is diverse, with specific alleles recognizing different strains of *Hpa* (Steinbrenner *et al.* 2012). The RPP NLRs responsible for hybrid necrosis show signs of diversifying selection in their LRR domains, potentially indicating they are locked in arms race with rapidly evolving *Hpa* effector

proteins. Furthermore, domain-swapping experiments show that amino-acid residues in the LRR domain are required for hybrid necrosis, suggesting that adaptation to *Hpa* indirectly results in hybrid incompatibility. Although interacting loci have been identified, it is currently unknown how they directly or indirectly interact with NLR to induce hybrid necrosis.

Although incompatibility genes can show signs of selection, they do not always have a clear role in promoting ecological adaptation (Tao *et al.* 2001; Ferree and Barbash 2009; Phadnis and Orr 2009; Seidel *et al.* 2011). For example, one emerging theme is that incompatibility genes are generated by intragenomic co-evolutionary arms races between selfish genetic elements and unlinked suppressors. For example, certain classes of selfish genetic elements act during spermatogenesis, increasing their chances of being inherited in offspring by poisoning the 50% of sperm that do not carry them. Because these selfish elements increase their frequency in a population by biasing their transmission rather than through organismal selection, they often carry mutations that are deleterious to organismal fitness. Therefore, unlinked loci that suppress the selfish elements biased transmission will restore organismal fitness and come to fixation. Repeated bouts of co-evolution between selfish genetic elements and their suppressors are thought to result in hybrid incompatibility when a normally suppressed selfish element is introduced into a naïve heterotypic genetic background. Consistent with this model for the evolution of hybrid incompatibility, several genetic loci have been mapped that cause both hybrid dysfunction and biased inheritance (Tao *et al.* 2001; Phadnis and Orr 2009; Seidel *et al.* 2011; Zanders *et al.* 2014). For example, crosses between two subspecies of *Drosophila pseudoobscura* produces sterile male hybrids. Interestingly, hybrid males (XY) become weakly fertile as they age, but almost all of their functional sperm carry the X-chromosome. It was found that both hybrid male sterility and the biased inheritance of the X-chromosome requires

Overdrive, a rapidly evolving DNA binding protein on the X-chromosome. It is hypothesized that *Overdrive* is a selfish genetic element that is normally suppressed in its native genomic background, but poisons Y-chromosome bearing sperm in hybrids. The poisoning mechanism may be improperly regulated, effecting all sperm early in the hybrid male's life and resulting in sterility. Therefore, there is evidence that non-adaptive change promotes the evolution of reproductive barriers and by extension, speciation.

There are currently only a handful of known incompatibility genes from a limited number of genera. Additional studies from a wider range of taxa are needed to gain a better understanding of the evolutionary forces that drive speciation. Despite the paucity of molecularly identified incompatibility genes, the segregation of deleterious phenotypes in a number of interspecific hybridizations indicates that incompatibilities between cytoplasmic and nuclear genomes occur frequently (Ellison and Burton 2008; Ellison *et al.* 2008; Sambatti *et al.* 2008; Arnqvist *et al.* 2010; Ross *et al.* 2011; Aalto *et al.* 2013). Furthermore, several studies have definitively mapped these incompatibility loci to the mitochondrial genome and nuclear genes with mitochondrial functions (Lee *et al.* 2008; Chou *et al.* 2010; Luo *et al.* 2013; Meiklejohn *et al.* 2013; Huang *et al.* 2015). Aerobic eukaryotic organisms rely on mitochondria to generate energy required for diverse biological processes. The mitochondrial genome encodes a small fraction of the mitochondrial proteins, all of which are components of the electron transport chain. Nuclear genes encode the majority of mitochondrial proteins, including components of the electron transport chain that physically interact with those encoded by the mitochondrial genome. Furthermore, the nuclear genome also encodes proteins that are required for the proper replication, transcription, and translation of mtDNA (Gustafsson *et al.* 2016). Given the interdependence of the nuclear and mitochondrial genomes, they are expected to coevolve by the

accumulation of compatible mutations that maintain mitochondrial function. By extension, distinct lineages that undergo unique mitochondrial– nuclear coevolution may be incompatible and result in mitochondrial dysfunction.

One especially common and well-studied form of mitochondrial-nuclear incompatibility results in reduced or abolished pollen production in hermaphroditic flowering plants (i.e. cytoplasmic male sterility, or CMS). CMS is a widespread phenomenon, with 140 species across 47 genera being affected (Laser and Lersten 1972; Burt and Trivers 2006). CMS often occurs in the hybrid offspring of different species or different populations of the same species. It is hypothesized that interspecific CMS is largely the result of intraspecific molecular arms races between selfish mitochondrial DNA and nuclearly encoded suppressors. The mitochondrial genome in flowering plants is only inherited through ovules. Therefore, mtDNA can increase its fitness by inhibiting the development of male reproductive tissues, diverting resources that would have been used in pollen production towards ovule production. However, this reproductive strategy is at odds with the nuclear genome, which is inherited equally through ovules and pollen. Therefore, as a CMS inducing mitochondrial genome increases in frequency, there will be selection for nuclear loci that suppress CMS and restore male fertility. Because male fertility is restored, the arms race between the two genomes is now cryptic and only revealed when the selfish mitochondrial genome is combined with a naïve nuclear genome through hybridization.

In many cases of CMS, the mtDNA component of the incompatibility is a novel open-reading frame generated by recombination of pre-existing genes encoded by the mitochondrial genome. There are few known examples of how these novel ORFs induce male sterility. Studies of CMS in wild rice show that a novel mitochondrially encoded ORF (*WA352*) is expressed in all

plant tissues, but the mature protein is only found in male reproductive tissues (Luo *et al.* 2013). WA352 binds to and likely inhibits the function of COX-11, a component of complex IV of the electron transport chain. There is an associated increase of reactive oxygen species and leakage of cytochrome-c into the cytoplasm, which likely induces the premature programmed cell death of the tapetum, the anther tissue that provides nutrients to developing pollen.

The nuclear suppressors of CMS are usually pentatricopeptide repeat (PPR) motif containing genes. The PPR motif binds to RNAs, and PPR genes are often transported into the mitochondria and chloroplasts where they regulate various steps of RNA maturation and post-transcriptional gene expression (Gorchs Rovira and Smith 2019). PPR genes typically restore male fertility by preventing the maturation or translation of CMS inducing novel ORF transcripts (Gaborieau *et al.* 2016). Consistent with their involvement in molecular arms races with selfish mtDNA, PPR genes across various species of hermaphroditic plants show signs of positive selection (Fujii *et al.* 2011). CMS is another convincing example of non-adaptive evolution mediating hybrid incompatibility and by extension, speciation.

Several other theories have been proposed to explain what drives the evolution of other known mitochondrial-nuclear genetic incompatibilities, including adaptation to different carbon sources (Lee *et al.* 2008), and the accumulation of deleterious mtDNA mutations and the evolution of compensatory nuclear variants that rescue mitochondrial function (Rand *et al.* 2004; Oliveira *et al.* 2008; Osada and Akashi 2012). However, given the scarcity of molecularly identified cases of mitochondrial–nuclear incompatibilities, additional studies are required to form more complete theories regarding the forces that drive their evolution.

Some studies on the genetic basis of hybrid incompatibility have focused on strong postzygotic reproductive barriers between well-defined species, and show that many genetic

variants contribute to dysfunction of hybrids (Coyne and Orr 1998). These studies are valuable, but it is difficult to determine the dynamics of the accumulation of such variants or their relative roles in initiating speciation. For example, theoretical work indicates that the number of genetic incompatibilities increases greater than linearly with the number of genetic differences between two lineages (Orr 1995). Therefore, a small number of genetic incompatibilities may initially reduce gene flow and promote genetic divergence between populations, whereas others evolve after strong reproductive barriers have already been established. Given this, studies of incomplete postzygotic barriers between young species or divergent populations within species are essential to understand the evolutionary forces that initiate speciation.

The nematodes of the *Caenorhabditis* genus are ideally suited for the genetic study of intrinsic post-zygotic reproductive barriers because they are easily crossed, have a short generation time of three days and because the mechanisms underlying *C. elegans* development have been heavily studied. However, previous studies of hybrid incompatibility within the genus were restricted by the limited number of known species and wild isolates. The recent discovery that *Caenorhabditis* "soil nematodes" are actually more commonly found in rotten fruit than soil has led to a continuously expanding number of wild-isolates of known and new species (Kiontke *et al.* 2011), greatly increasing the number of crosses in which intra and interspecies incompatibilities can be studied. Given this, the *Caenorhabditis* genus has great potential to aid in the rapid discovery of incompatibility genes and therefore expand our knowledge of the evolutionary forces that drive speciation.

Here we report incompatibility between the cytoplasmic and nuclear genomes of two distinct wild isolates of the male-female nematode *Caenorhabditis nouraguensis*. Cytoplasmic-nuclear incompatibility is not specific to these two strains, but is also observed upon

hybridization of other distinct wild isolates of *C. nouraguensis*, indicating that this is a naturally widespread reproductive barrier within the species. This cytoplasmic-nuclear incompatibility may provide an excellent opportunity for a detailed study of mitochondrial-nuclear incompatibility, the forces that drive the coevolution of these genomes, and their possible role in speciation.

1.2 THE EVOLUTION OF ASEXUALITY FROM A SEXUAL ANCESTOR

Theory predicts that facultative sex, the ability to undergo both asexual and sexual reproduction, is the optimal reproductive strategy (Green and Noakes 1995; D'Souza and Michiels 2010; Stelzer and Lehtonen 2016; Burke and Bonduriansky 2017). Advantages of asexuality include an immediate two-fold enhancement of fitness and an enhanced ability to disperse geographically through not requiring a mate for reproduction (Maynard Smith 1971; Gibson *et al.* 2017). Advantages of sexual reproduction include the production of genotypic diversity that could be used to adapt to a changing environment and the ability to purge deleterious alleles through recombination (Felsenstein 1976). Most unicellular eukaryotes undergo facultative sex, taking advantage of their ability to switch between these two reproductive strategies as conditions dictate (Kassir *et al.* 1988; Dacks and Roger 1999). And yet, despite the predicted advantages of facultative sex, most animal species are obligately sexual, suggesting that there must be strong selective pressures to prevent the origin or persistence of asexuality from an obligately sexual ancestor (Burke and Bonduriansky 2017). A better understanding of these selective pressures requires understanding how asexuality evolves from a sexual ancestor. However, very few such transitions are known, and even fewer occur in genetically tractable organisms.

Here we focus on one key aspect of how asexuality evolves from sexually reproducing organisms: diploid maternal inheritance. Sexual reproduction requires that diploid females and males generate haploid eggs and sperm, which then fuse to produce the next generation of diploid offspring. By contrast, asexual females produce diploid eggs that either develop independently of sperm fertilization, known as parthenogenesis (Mirzaghaderi and Hörandl 2016), or require fertilization but do not inherit the paternal genome, known as gynogenesis (Beukeboom and Vrijenhoek 1998). Thus, understanding how egg production can be modified to result in diploid maternal inheritance provides insight into understanding the origins of asexuality.

There are several known mechanisms of generating diploid eggs in asexual species. In apomixis, eggs are produced by mitotic divisions rather than meiosis, resulting in offspring that are clones of their mother. In contrast, automixis maintains the first steps of meiosis, but generates diploid eggs either by combining two haploid products of meiosis or duplicating one of the four meiotic products (Mirzaghaderi and Hörandl 2016). Apomictic (mitotic) parthenogenesis maintains heterozygosity across the entire genome and prevents inbreeding depression, whereas automictic parthenogenesis via duplication of a single meiotic product results in genome-wide homozygosity and inbreeding depression. Thus, different mechanisms of generating diploid eggs to establish new asexual lineages can have a range of genetic consequences that influence their success (Engelstädter 2008).

Interestingly, in some predominantly sexual invertebrate and vertebrate species, females can reproduce asexually at a low frequency (Groot *et al.* 2003; Watts *et al.* 2006; Eisman and Kaufman 2007; Schwander *et al.* 2010; Markow 2013; Chang *et al.* 2014; Fields *et al.* 2015; van der Kooi and Schwander 2015). Supporting the hypothesis that rare asexuality is an intermediate

step in the transition between sexual and asexual reproduction (Schwander *et al.* 2010; van der Kooi and Schwander 2015), genetic studies indicate that rare asexual offspring inherit a diploid maternal genome by mechanisms like that of obligately asexual species. For example, detailed genetic and cytological studies in several sexual *Drosophila* species show that their rare parthenogenetic offspring inherit a diploid maternal genome by automixis (Templeton *et al.* 1976; Fuyama 1986a, 1986b; Chang *et al.* 2014). Furthermore, the frequency of rare parthenogenesis in some *Drosophila* species can be increased through experimental laboratory selection (Stalker 1954; Fuyama 1986b; Markow 2013), suggesting that rare asexuality could be a transition step en route to obligate asexuality. However, other than *Drosophila*, most sexual species that produce rare asexual offspring are not model organisms. As a result, genetic studies in these non-model species have been typically limited to a handful of genetic markers. Although highly informative, these studies have not revealed sufficient mechanistic insights into the basis of diploid maternal inheritance.

Here we report that a cross between two sexual *Caenorhabditis* nematode species results in rare viable progeny that are fertile or sterile. Fertile offspring are generated by gynogenesis (sperm-dependent parthenogenesis), whereas sterile offspring are hybrids. Interestingly, both fertile and sterile offspring inherit a diploid maternal genome by automixis, specifically the inheritance of two homologous chromatids. We also find that intraspecific crosses within one of the species, *C. nouraguensis*, can result in a low frequency of gynogenetic reproduction. The low frequency of asexuality seen in this species may be indicative of a nascent asexual lineage.

Chapter 2. CYTOPLASMIC-NUCLEAR INCOMPATIBILITY BETWEEN WILD-ISOLATES OF *CAENORHABDITIS NOURAGUENSIS*[1]

2.1 INTRODUCTION

How species arise is a fundamental question in biology. Species can be defined as populations of interbreeding individuals that are reproductively isolated from other such populations. Therefore, understanding how reproductive barriers evolve between populations is essential for understanding the process of speciation. Hybrid incompatibility (for example, hybrid sterility or lethality) is a common and strong reproductive barrier in nature. Here we report a lethal incompatibility between two wild isolates of the nematode *Caenorhabditis nouraguensis*. Hybrid inviability results from the incompatibility between a maternally inherited cytoplasmic factor from each strain and a recessive nuclear locus from the other. We have excluded the possibility that maternally inherited endosymbiotic bacteria cause the incompatibility by treating both strains with tetracycline and show that hybrid death is unaffected. Furthermore, cytoplasmic-nuclear incompatibility commonly occurs between other wild isolates, indicating that this is a significant reproductive barrier within *C. nouraguensis*. We hypothesize that the maternally inherited cytoplasmic factor is the mitochondrial genome and that mitochondrial dysfunction underlies hybrid death. This system has the potential to shed light on the dynamics of divergent mitochondrial-nuclear coevolution and its role in promoting speciation.

[1] This is chapter is closely adapted from (Lamelza and Ailion 2017)

2.2 RESULTS

2.2.1 *Two strains of C. nouraguensis exhibit F2 hybrid breakdown*

Two strains of *C. nouraguensis*, JU1825 and NIC59, were derived from single gravid females that were isolated approximately 112 kilometers apart in French Guiana (Kiontke *et al.* 2011). Both of these strains were designated as *C. nouraguensis* based on having highly similar ITS2 rDNA sequences (a good species barcode within the *Caenorhabditis* genus), and because they produced many viable F1 offspring when crossed (Kiontke *et al.* 2011; Félix *et al.* 2014). We found that both strains produce high numbers of viable progeny in intra-strain crosses. We also confirmed the previous finding of F1 hybrid viability by crossing NIC59 females to JU1825 males, and vice versa, showing that the F1 hybrids resulting from these inter-strain crosses exhibit levels of viability comparable to those seen in intra-strain crosses (Figure 2.1A).

However, not all reproductive barriers act in the F1 generation. There are many cases of F2 hybrid breakdown, in which reduction of hybrid fitness is seen in the F2 generation due to recessive incompatibility loci (Masly *et al.* 2006; Bikard *et al.* 2009; Dey *et al.* 2012, 2014; Stelkens *et al.* 2015). To test for F2 hybrid inviability, we mated hybrid F1 siblings derived from either JU1825 female x NIC59 male crosses, or from NIC59 female x JU1825 male crosses, and assayed the F2 generation for reductions in fitness. These F1 hybrids are referred to as "(J); N/J" and "(N); N/J" respectively, where the genotype is designated by the following nomenclature: (cytoplasmic genotype); nuclear genotype. The cytoplasmic genotype indicates genetic elements that are inherited only maternally, such as the mitochondrial genome. We found that both types of F1 sibling crosses resulted in a significant decrease in the percentage of viable progeny, with on average only 71% and 63% of F2 embryos maturing to the L4 or young adult stage (Figure 2.1A). These results indicate that there are divergent genomic loci between NIC59 and JU1825

that cause inviability only when they become homozygous in F2 hybrids. Additionally, there is no difference in sex-specific mortality in hybrids in comparison to intra-strain crosses (Figure 2.1B), implying that these loci are autosomally linked, as we show later.

2.2.2 *Incompatibilities between cytoplasmic and nuclear genomes cause F2 inviability*

To further understand the genetic architecture of hybrid breakdown between JU1825 and NIC59, we tested whether maternally or paternally inherited factors are required for F2 inviability. We reasoned that backcrossing F1 females to parental males would test whether maternal factors are required for reduced hybrid fitness, while backcrossing F1 males to parental females would test whether paternal factors are required. For example, backcrossing F1 hybrid females to JU1825 males will result in an F2 population with a 50% chance of being heterozygous (NIC59/JU1825) and a 50% chance of being homozygous (JU1825/JU1825) for any given autosomal locus. Therefore, this cross will test for maternally deposited NIC59 factors that are incompatible with homozygous JU1825 autosomal loci. The same logic can be applied to crosses of F1 hybrid males to parental strain females.

All backcrosses of F1 hybrid males to parental strain females resulted in levels of F2 viability similar to those observed in parental strains. Therefore, paternal factors do not have a major effect on F2 inviability (Figure 2.2A). Only two crosses consistently resulted in significantly reduced viability. The first is when (N); N/J F1 females were crossed to JU1825 males, with on average only 36% of F2 hybrids maturing to the L4 or young adult stage. This cross implies that there are maternally derived NIC59 factors distributed to F2 embryos, and these factors are incompatible with recessive JU1825 nuclear loci. The second is when (J); N/J F1 females are crossed to NIC59 males, with on average only 52% of the F2 hybrids maturing to

the L4 or young adult stage (Figure 2.2B). This cross implies that there are also maternally derived JU1825 factors distributed to F2 embryos, and these factors are incompatible with recessive NIC59 nuclear loci. The viability of (J); N/J F1 female x JU1825 male crosses can also be significantly reduced in comparison to intra-strain crosses, but varies within and between experiments (Figure 2.7).

The F1 female backcross experiments show that almost identical crosses, which differ only in the cytoplasmic genotype of the F1 female, have significantly different rates of F2 viability. For instance, (N); N/J F1 female x JU1825 male crosses consistently have significantly lower F2 viability than (J); N/J F1 female x JU1825 male crosses (Figure 2.2, Figure 2.7). Similarly, (J); N/J F1 female x NIC59 male crosses consistently have significantly lower F2 viability than (N); N/J F1 female x NIC59 male crosses (Figure 2.2B). The F1 hybrid females in these pairs of crosses are expected to be genotypically identical at all nuclear loci, suggesting that something other than the F1 nuclear genome encodes maternal factors that lead to F2 inviability.

One model to explain these backcrosses is that the mitochondrial genome is the maternally inherited factor that is incompatible with recessive nuclear loci in the F2 generation. For example, all F2 progeny from (N); N/J F1 female x JU1825 male crosses will inherit only NIC59 mtDNA, which may be incompatible with nuclear loci homozygous (or hemizygous) for JU1825 alleles, resulting in inviability (Figure 2.6A). In comparison, all F2 progeny from (J); N/J F1 female x JU1825 male crosses will inherit only JU1825 mtDNA, which should be compatible with the JU1825 nuclear genome and therefore not result in the same inviability. The same logic can be applied to the (J); N/J F1 female x NIC59 male and (N); N/J F1 female x NIC59 male crosses. We hypothesize that F2 inviability is the result of two mitochondrial-

nuclear incompatibilities, one between the NIC59 mitochondrial genome and recessive JU1825 nuclear loci, and another between the JU1825 mitochondrial genome and recessive NIC59 nuclear loci.

2.2.3 *The nuclear incompatibility loci are linked to autosomes*

Nematodes commonly have an XX (female) and XO (male) sex determining mechanism (Pires-daSilva 2007). The F1 hybrid female backcross experiments reveal that there is no difference in sex-specific mortality in hybrids in comparison to intra-strain crosses (Figure 2.2C). However, given the expected genotypes of their F2 populations, these backcrosses on their own do not allow us to determine whether the nuclear incompatibility loci are autosomally or X-linked. In the previous section, we concluded that the inviability of the F2 progeny derived from (N); N/J F1 female x JU1825 male crosses is the result of a genetic incompatibility between the NIC59 cytoplasmic genome and nuclear loci homozygous (or hemizygous) for JU1825 alleles. If this is true, it is reasonable to assume that the same genetic incompatibility occurs in (N); N/J F1 female x (N); N/J F1 male crosses (Figure 2.1A). In this F1 sibling cross, if the JU1825 nuclear incompatibility locus were autosomally linked, both sexes would suffer equal rates of inviability. However, if the nuclear incompatibility locus were linked to the X-chromosome, then we would expect a significant decrease in the proportion of viable males in comparison to intra-strain crosses (Figure 2.8). However, we observe no significant difference in the proportion of viable males for the (N); N/J F1 female x (N); N/J F1 male cross (Figure 2.1B). Therefore, given the data from the F1 female backcrosses and the F1 sibling crosses, we conclude that the JU1825 nuclear incompatibility locus is autosomally linked. A similar line of reasoning indicates that the NIC59 nuclear incompatibility locus is also autosomally linked.

2.2.4 *Endosymbiotic bacteria do not cause hybrid inviability*

We hypothesize that mitochondrial genomes are responsible for the cytoplasmic component of the hybrid incompatibility between NIC59 and JU1825. However, we also considered whether endosymbiotic bacteria of the *Rickettsiales* order could be involved. Within this order, bacteria of the *Wolbachia* genus are known to infect certain species of nematodes, and are transmitted to host progeny through female gametes (Werren *et al.* 2008). Furthermore, hybrid lethality in inter-strain and interspecies crosses is sometimes caused by infection with divergent *Wolbachia* strains (Bourtzis *et al.* 1996; Bordenstein *et al.* 2001). However, we failed to detect conserved genes typically used to genotype diverse strains of *Wolbachia* in either JU1825 or NIC59 using PCR with degenerate primers (Figure 2.9A). Additionally, treatment of both strains with tetracycline for nine generations failed to rescue hybrid inviability (Figure 2.9B). Endosymbiotic bacteria within the *Rickettsiales* order are typically susceptible to tetracycline (McOrist 2000; Darby *et al.* 2015). Thus, endosymbiotic bacteria are unlikely to cause the reproductive barrier between NIC59 and JU1825.

2.2.5 *Cytoplasmic-nuclear incompatibility is common within C. nouraguensis*

We hybridized additional wild isolates (Figure 2.3A) to determine whether cytoplasmic-nuclear incompatibilities represent a common reproductive barrier within *C. nouraguensis,* or whether they are an unusual phenotype only observed in hybridizations between NIC59 and JU1825. Specifically, we tested the compatibility of four cytoplasmic genotypes with seven nuclear genotypes. To test for an incompatibility between one strain's cytoplasm and another strain's nuclear genome, we again compared the viabilities of backcrosses that differ only in the F1 hybrid female's cytoplasmic genotype (Figure 2.3B). Specifically, we compared the viability

of the backcross that combines heterotypic cytoplasmic and nuclear genotypes to the viability of the backcross that combines homotypic cytoplasmic and nuclear genotypes. We calculated the relative viability of the two crosses (heterotypic combination/homotypic combination), and tested for statistically significant differences (see Materials and Methods). Using the same logic as for our JU1825 x NIC59 crosses, we reasoned that lower viability of the heterotypic cytoplasmic-nuclear combination in comparison to the homotypic cytoplasmic-nuclear combination indicates a cytoplasmic-nuclear incompatibility. Three or four biological replicates were performed for each cytoplasmic-nuclear combination.

Of the 74 cytoplasmic-nuclear tests performed, 50 (67%) exhibited significant incompatibilities (Figure 2.3C). Additionally, each cytoplasmic genotype was consistently incompatible with at least one heterotypic nuclear genotype (that is, all replicates for a particular cytoplasmic-nuclear combination indicate a significant incompatibility). However, there are a number of cytoplasmic-nuclear combinations whose replicates are inconsistent with one another (that is, some replicates indicate a significant incompatibility while others do not) (Figure 2.3D and Figure 2.10). This may indicate that the genetic loci required for hybrid inviability are not fixed between the strains, but rather are polymorphisms segregating within each strain (Cutter 2012; Kozlowska *et al.* 2012; Corbett-Detig *et al.* 2013), consistent with the fact that none of these strains have been formally inbred. Regardless, given their common occurrence in hybridizations between strains of *C. nouraguensis,* we hypothesize that cytoplasmic-nuclear incompatibilities are a significant reproductive barrier within the species.

We generated a heat map to help visualize the median relative viability for each cytoplasmic-nuclear combination (Figure 2.3D). Strikingly, the NIC59 cytoplasmic genotype exhibits a distinct response to hybridization, being strongly incompatible (that is, having a low

median relative viability) with all of the nuclear genotypes tested. By comparison, the other cytoplasmic genotypes can be relatively compatible with some heterotypic nuclear genotypes or exhibit incompatibilities that are typically weaker than those involving the NIC59 cytoplasmic genotype. Specifically, incompatibilities involving the JU1837 or JU1854 cytoplasmic genotypes have significantly higher relative viability (median=0.72 and 0.71, respectively) in comparison to incompatibilities with the NIC59 cytoplasmic genotype (median=0.45) (Figure 2.3C). Incompatibilities involving the JU1825 cytoplasm exhibit an intermediate level of relative viability (median=0.64) that is statistically indistinguishable from the other cytoplasmic genotypes (P=0.057 in comparison to NIC59; P=1.0 in comparison to both JU1837 and JU1854). Although there is a correlation between the severity of cytoplasmic-nuclear incompatibility and geographic location of the strains hybridized (Figure 2.3A), too few strains were tested to conclude that the incompatibility studied here has already led to reproductive isolation of these allopatric populations. However, it is clear that the NIC59 cytoplasmic genotype is distinct in terms of the nuclear genotypes it is incompatible with and how severe those incompatibilities are.

2.2.6 *A single BDM incompatibility between a NIC59 cytoplasmic locus and a JU1825 nuclear locus causes embryonic lethality*

As previously discussed, the backcross that combines the NIC59 cytoplasmic genotype with JU1825 nuclear genotype (that is, (N); N/J F1 female x JU1825 male, Figure 2.2B) results in only ~36% of F2 offspring maturing to the L4 or young adult stage. A more detailed characterization of F2 inviability shows that ~50% of F2 offspring fail to complete embryogenesis (Figure 2.4A). Of the remaining half that complete embryogenesis, ~33% fail to mature to the L4 or young adult stage (data not shown). In comparison, (J); N/J F1 female x

JU1825 male crosses result in low levels of embryonic lethality, similar to parental crosses. These data are consistent with F2 embryonic lethality resulting from a single BDM incompatibility between a NIC59 cytoplasmic locus and a single homozygous JU1825 autosomal locus.

To test the hypothesis of a single BDM incompatibility, we crossed F1 (N); N/J females to JU1825 males, then crossed the viable F2 females to JU1825 males and assayed F3 viability. Under this hypothesis, the surviving F2 females are expected to have inherited NIC59 mtDNA and be heterozygous (that is, JU1825/NIC59) at the JU1825 nuclear incompatibility locus (Figure 2.6A). Therefore, crossing these F2 females to JU1825 males should also result in ~50% embryonic lethality in the F3 generation. This pattern should also be true for additional backcross generations (F4, F5 etc.). Thus, we generated 15 independent backcross lineages, each consisting of matings between single surviving hybrid females and JU1825 males, and monitored each lineage's viability for four backcross generations. Indeed, the approximately 50% embryonic lethality observed in the F2 generation is also observed in the subsequent backcross generations in all lineages (Figure 2.4B). These results are consistent with the hypothesis that embryonic lethality is the result of a simple two-locus BDM incompatibility between a purely maternally inherited cytoplasmic NIC59 locus and a single nuclear locus homozygous for JU1825 alleles. We hypothesize that the post-embryonic inviability may be a genetically separable phenotype.

2.2.7 *The JU1825 cytoplasm appears to be heteroplasmic*

As previously discussed, the backcross that combines the JU1825 cytoplasmic genotype with the NIC59 nuclear genotype (that is, (J); N/J F1 female x NIC59 male crosses) results in

~50% F2 viability on average (Figure 2.2B). Thus, the total F2 inviability could be the result of a single BDM incompatibility between a JU1825 cytoplasmic locus and a single autosomal locus homozygous for NIC59 alleles.

To test this hypothesis, we generated 14-15 independent backcross lineages, each consisting of matings between single surviving (J); N/J hybrid females and NIC59 males, and monitored each lineage's viability for four backcross generations. To our surprise, though some lineages continued to exhibit low levels of viability similar to the F2 generation average (~50%), others began to exhibit and maintain significantly increased viability for multiple backcross generations (Figure 2.5A). For example, in this particular experiment we found that in the F2 generation a majority of lineages (13/15) had a total viability ranging from 18-50%, while only two exhibited higher viability (68% and 85%). However, by the F5 backcross generation, we found that of the fourteen remaining lineages only four exhibited 50% viability or less. Strikingly, by the F5 generation, 5/14 backcross lineages exhibited nearly 100% viability.

The rescue of hybrid inviability for some lineages via several generations of backcrossing is peculiar. One hypothesis to explain this phenomenon is that the JU1825 cytoplasmic or NIC59 nuclear incompatibility loci are not fixed within their respective strains, but rather are segregating polymorphisms (Cutter 2012; Kozlowska *et al.* 2012; Corbett-Detig *et al.* 2013). As a specific example, the JU1825 cytoplasmic incompatibility locus could be heteroplasmic for alleles that are either incompatible or compatible with the NIC59 nuclear genome. The mitochondrial genome is present at a high copy number within a single cell, and it is thought that individual mtDNAs are randomly replicated and segregated to daughter cells during cell division. Studies on the inheritance of various mtDNA heteroplasmies show that their frequency amongst siblings from the same mother can be highly variable due to the random sampling of

mtDNAs and genetic bottlenecks during female germline development (Wallace and Chalkia 2013; Gitschlag *et al.* 2016). Therefore, it is possible that a NIC59-compatible cytoplasmic allele has increased in frequency in some backcross lineages and rescued inviability.

To gain a better understanding of the genetic composition of the JU1825 cytoplasm, we also monitored the viability of (J); N/J female x JU1825 male lineages over four backcross generations. Because this cross combines homotypic JU1825 cytoplasmic and JU1825 nuclear genotypes, we originally predicted that the relatively high rates of F2 viability would persist or possibly increase with additional backcross generations. However, we instead observed that some backcross lineages showed a striking decrease in viability after the F2 generation (Figure 2.5B). For example, in this particular experiment, lineages in the F2 generation exhibited a uniform distribution of viability, with an average of 74%. By the F5 generation we find two distinct populations of lineages, those with a high viability ranging from 85-96% (6/14 lineages) and those with low viability ranging from 29-55% (8/14 lineages) (Figure 2.5B). The latter population has an average viability of 39%, which is similar to that observed in (N); N/J F1 female x JU1825 male crosses (~36%, Figure 2.2B), indicating that although these lineages inherited their cytoplasm from JU1825 mothers, they now seem to exhibit low levels of viability similar to those observed in the NIC59 cytoplasmic–JU1825 nuclear incompatibility. One hypothesis to explain these data is that the JU1825 cytoplasm harbors a NIC59-like allele which at a certain threshold frequency can mimic the NIC59 cytoplasmic-JU1825 nuclear incompatibility in certain (J); N/J F1 female x JU1825 male backcross lineages.

In support of this hypothesis, the rate of embryonic lethality for some (J); N/J female x JU1825 male backcross lineages also increases to levels observed in the NIC59 cytoplasmic–JU1825 nuclear incompatibility (that is, 50%) and can be stably inherited for several backcross

generations (Figure 2.5C). Specifically, most lineages (12/14) in the F2 generation exhibited only 0-19% embryonic lethality, whereas two lineages exhibited higher rates (38 and 47%). However, by the F5 backcross generation, only about half of the lineages (6/14) exhibited 0-8% embryonic lethality, whereas 8/14 lineages exhibited 35-65% embryonic lethality. Taken together, the results from the two backcross experiments are consistent with the hypothesis that the JU1825 cytoplasm is heteroplasmic and harbors both JU1825-like and NIC59-like incompatibility loci (Figure 2.6B and C).

2.3 DISCUSSION

We discovered a lethal cytoplasmic-nuclear incompatibility between two wild isolates of *C. nouraguensis*, JU1825 and NIC59, and find that such incompatibilities may be widespread between other wild isolates within the species. We propose that the mitochondrial genome is the most likely candidate for harboring the cytoplasmic incompatibility factor(s) and further propose that the JU1825 cytoplasm is heteroplasmic and harbors both JU1825-like and NIC59-like incompatibility loci. We show that maternally inherited endosymbiotic bacteria are probably not the cause of hybrid inviability. It remains possible that incompatibility is caused by other cytoplasmically inherited factors (such as maternally inherited small RNAs), or by maternal inheritance of epigenetic marks across several generations.

In eukaryotes, the mitochondrial genome typically contains a very small fraction of the gene content of a cell, yet it seems to be involved in a disproportionate number of genetic incompatibilities across a diverse range of taxa (Rand *et al.* 2004; Burton and Barreto 2012). However, there are relatively few cases in which incompatibility loci have been definitively mapped to the mitochondrial genome, and therefore a larger sample is required to better

understand what drives the evolution of mitochondrial-nuclear incompatibility. Additionally, all of the molecularly identified cases of mitochondrial-nuclear incompatibility have been found between species rather than within species (Lee *et al.* 2008; Chou *et al.* 2010; Luo *et al.* 2013; Meiklejohn *et al.* 2013; Ma *et al.* 2016). Some of these inter-species hybridizations harbor additional genetic incompatibilities or chromosomal rearrangements that cause inviability and sterility (Hunter *et al.* 1996; Fischer *et al.* 2000; Brideau *et al.* 2006; Ferree and Barbash 2009; Mihola *et al.* 2009; Davies *et al.* 2016), making it difficult to discern whether mitochondrial-nuclear incompatibility was instrumental in initiating speciation or evolved after strong reproductive isolation occurred. The incompatibility we describe here provides an excellent opportunity to study the evolutionary genetics and cell biology of incipient speciation as well as mitochondrial-nuclear incompatibility. The ease of breeding, large brood sizes, and short generation time of *C. nouraguensis* should facilitate the mapping and identification of the genes that contribute to hybrid inviability.

2.3.1 *Cytoplasmic-nuclear incompatibility: both sexes are equally inviable*

J.B.S Haldane noted that the heterogametic sex more often suffers from inviability or sterility in inter-species hybridizations than the homogametic sex (Delph and Demuth 2016). This rule holds for the handful of recently studied inter-species hybridizations in *Caenorhabditis* (Baird 2002; Woodruff *et al.* 2010; Kiontke *et al.* 2011; Dey *et al.* 2012, 2014; Kozlowska *et al.* 2012; Ragavapuram *et al.* 2016). However, it is not known whether Haldane's rule also generally applies to intra-species hybridizations. Interestingly, some intra-species incompatibilities in *Caenorhabditis* affect both sexes equally (Seidel *et al.* 2008, 2011; Huang *et al.* 2014).

The lethal cytoplasmic-nuclear incompatibility we identified between the NIC59 and JU1825 wild isolates of *C. nouraguensis* also affects females and males equally, suggesting that the two sexes share the same disrupted developmental pathway(s). However, we have not carefully studied other aspects of sex-specific fitness, such as female and male F2 hybrid fertility. Because the mitochondrial genome is inherited only through females, theory predicts that evolution will lead to the accumulation of mtDNA variants that are neutral or increase female fitness, but that are neutral or possibly deleterious to male fitness (Gemmell *et al.* 2004, Patel *et al.* 2016). Thus, male-specific functions may be more adversely affected during the hybridization of heterotypic mitochondrial and nuclear genomes. This is indeed the case for some known mitochondrial-nuclear incompatibilities. For example, when swapping the mitochondrial genomes between mouse subspecies via pronuclear transfer, one mitochondrial-nuclear combination resulted in reduced male fertility whereas females had relatively normal fertility (Ma *et al.* 2016). Therefore, further studies of *C. nouraguensis* hybrid male fertility will be required to more fully address whether this system follows Haldane's rule, as well as to determine whether there are male-specific mitochondrial-nuclear incompatibilities.

2.3.2 *Symmetric cytoplasmic-nuclear incompatibilities in C. nouraguensis*

Reciprocal interspecific crosses often show differences in the viability or fertility of hybrids. This asymmetry in hybrid fitness (termed "Darwin's corollary" to Haldane's rule) has been theorized to be the result of uniparentally inherited factors from one species (such as maternal RNAs, sex chromosomes, or cytoplasmically inherited genomes), being incompatible with heterospecific loci of the other, but not vice versa (Turelli and Moyle 2007). Darwin's corollary is also seen in several hybridizations in the *Caenorhabditis* genus, probably due to X-

linked incompatibilities (Woodruff *et al.* 2010; Dey *et al.* 2012, 2014; Kozlowska *et al.* 2012; Ragavapuram *et al.* 2016).

Consistent with Darwin's corollary to Haldane's rule, most molecularly characterized BDM incompatibilities are asymmetric, in that only one of two divergent alleles at a locus is incompatible with heterospecific alleles at other loci (Brideau *et al.* 2006; Ferree and Barbash 2009). This is also true of the asymmetric mitochondrial-nuclear incompatibilities seen in *Saccharomyces* species hybridizations (Lee *et al.* 2008; Chou *et al.* 2010). For example, an intron of the *COX1* gene in the *Saccharomyces bayanus* mitochondrial genome fails to be correctly spliced by the nuclearly encoded *S. cerevisiae MRS1* gene, resulting in hybrid inviability. However, a similar incompatibility does not occur between *S. cerevisiae COX1* and *S. bayanus MRS1*. In our study, despite differences in severity, cytoplasmic-nuclear incompatibilities involving NIC59 appear to be symmetric (Figure 2.2A and Figure 2.3D). However, with our current data, we cannot determine whether the same or different genes cause hybrid inviability in the reciprocal crosses. Multiple distinct cytoplasmic-nuclear incompatibilities between these strains might be an indication of rapid divergent cytoplasmic-nuclear coevolution within the species.

2.3.3 *JU1825 heteroplasmy*

We hypothesize that the JU1825 cytoplasm is heteroplasmic and contains mitochondrial genomes that are both compatible (JU1825-like) and incompatible (NIC59-like) with the JU1825 nuclear incompatibility locus. If the JU1825 cytoplasm is naturally heteroplasmic, we predict the NIC59-like mtDNAs are kept at a low frequency within JU1825 by selection. This selection would be relaxed in (J); N/J F1 hybrids and the frequency of NIC59-like mtDNA could increase

beyond a certain threshold, reducing incompatibility in backcrosses to NIC59 males and increasing incompatibility in backcrosses to JU1825 males. However, another possibility is that NIC59-like mtDNA is introduced into F1 females by incomplete degradation and inheritance of paternal NIC59 mtDNA. Interestingly, evidence suggests that paternal mtDNA can be inherited when hybridizing different wild isolates of *Caenorhabditis briggsae* (Hicks *et al.* 2012; Chang *et al.* 2015; Ross *et al.* 2016).

The hypothesized heteroplasmy may explain the greater variance of F2 viability in crosses with (J); N/J F1 females in comparison to those with presumably homoplasmic (N); N/J F1 females. Stochastic segregation and genetic bottlenecking events from JU1825 mothers (or variable paternal leakage from NIC59 fathers) may result in F1 females with a wide range of frequencies of the NIC59-like cytoplasmic allele, and therefore a wide range of F2 viability when backcrossed to either NIC59 or JU1825 males. Such stochastic inheritance could explain why the degree of F2 viability of (J); N/J F1 female x JU1825 male backcrosses can also vary significantly from experiment to experiment (Figure 2.7).

2.3.4 *Caenorhabditis nematodes as models to study speciation*

The nematodes of the *Caenorhabditis* genus are currently emerging as a model system for the genetic study of hybrid incompatibility. Previous studies were restricted by the limited number of known species and wild isolates. However, the recent discovery that *Caenorhabditis* nematodes are found primarily in rotting fruits has led to a continuously expanding number of wild isolates of known and new species, greatly increasing the number of crosses in which intra and inter-species incompatibilities can be studied (Kiontke *et al.* 2011).

Studies of genetic incompatibilities between well-defined species often reveal that many genetic variants contribute to hybrid dysfunction, making it difficult to discern which initially decreased gene flow and which evolved after strong reproductive barriers had evolved. On the other hand, incomplete reproductive barriers between different populations of the same species may or may not be indicative of incipient speciation. Therefore, to understand the accumulation of post-zygotic isolating barriers, one would ideally monitor the same two divergent lineages throughout the entire speciation process (Seehausen *et al.* 2014). This is impractical for most multicellular organisms. An alternative method is to compare and contrast hybridizations with differing degrees of post-zygotic isolation across the species continuum, ranging from weak post-zygotic isolation within species to strong post-zygotic isolation between distinct species.

The *Caenorhabditis* genus has the potential to span such a continuum. Interestingly, both *C. briggsae* and *C. nouraguensis* appear to have intra-species cytoplasmic-nuclear incompatibilities (Ross *et al.* 2011; Chang *et al.* 2015). Although the exact genetic components of these incompatibilities have not been identified, these two cases add to an already large literature of cytoplasmic-nuclear incompatibilities, implying a role for divergent cytoplasmic-nuclear coevolution in driving speciation. Near the other end of the species continuum, hybridizations between the well-defined sister-species *C. briggsae* and *C. nigoni* produce a low degree of F1 embryonic lethality and either hybrid male sterility or inviability, depending on the cross direction (Woodruff *et al.* 2010; Kozlowska *et al.* 2012; Ragavapuram *et al.* 2016). In contrast to the relatively simple intra-species genetic incompatibilities in *C. nouraguensis, C. briggsae* and *C. elegans* (Seidel *et al.* 2008, 2011; Ross *et al.* 2011; Baird and Stonesifer 2012), a recent genome-wide introgression study revealed the presence of many distinct *C. briggsae* loci that are sufficient to cause hybrid dysfunction in an otherwise *C. nigoni* background (Bi *et al.*

2015). Future identification and comparison of genes required for hybrid inviability or sterility across the *Caenorhabditis* speciation continuum may give insight into the evolutionary forces that promote speciation.

2.4 Materials and Methods

2.4.1 *Strain isolation and maintenance*

All strains of *C. nouraguensis* used in this study were derived from single gravid females isolated in 2009 or 2011 from rotten fruit or flowers found in French Guiana (Kiontke *et al.* 2011; Félix *et al.* 2013, Christian Braendle (personal communication))**,** and have not been subjected to further inbreeding. Strains were kindly provided by Marie-Anne Félix ("JU" prefix) and Christian Braendle ("NIC" prefix). Strain stocks were stored at -80°C. Thawed strains were maintained at 25°C on standard NGM plates spread with a thin lawn of OP50 bacteria (Brenner 1974).

2.4.2 *Hybridizing JU1825 and NIC59*

To quantify inviability, we crossed one virgin L4 female and male, with 10-15 replicates for each cross. The edge of each plate was coated with a palmitic acid solution (10 mg/mL in 95% ethanol) and allowed to air dry, resulting in a physical barrier that helps prevent worms from leaving the plate's surface. The plates were placed at 25°C overnight, during which the worms matured to adulthood and began mating. The next day, each female-male couple was placed onto a new plate streaked with OP50 and rimmed with palmitic acid. Each couple was then allowed to mate and lay eggs for 5 hours at 25°C, and then were permanently removed. The embryos laid within those 5 hours were counted immediately. Approximately 17 hours later, we

counted the number of embryos that failed to hatch per plate. These unhatched embryos were scored as dead since *C. nouraguensis* embryogenesis is normally completed within 13 hours at 25°C (data not shown). We defined the percentage of embryonic lethality as the number of unhatched embryos divided by the total number of embryos laid. Approximately 20 hours later, we placed the plates at 4°C for an hour and then counted the number of healthy L4 larvae and young adults per plate. We defined the percentage of viable progeny as the total number of L4 larvae and young adults divided by the total number of embryos laid.

2.4.3 *Determining cytoplasmic-nuclear compatibility between various strains of C. nouraguensis*

The genotype of a strain is designated by the following nomenclature: (cytoplasmic genotype); nuclear genotype. The cytoplasmic genotype indicates genetic elements that are inherited only maternally, such as the mitochondrial genome. To test for an incompatibility between one strain's cytoplasm and another strain's nuclear genome, we compared the viabilities of backcrosses that differ only in the F1 hybrid female's cytoplasmic genotype (for example, (NIC59); NIC59/JU1837 F1 female x JU1837 male vs (JU1837); NIC59/JU1837 F1 female x JU1837 male, Figure 2.3B). We performed a Fisher's exact test to determine whether there were significant differences in the proportions of viable and inviable F2 progeny between the two types of crosses. We also calculated the relative viability of the two crosses (for example, the percent viability of the (NIC59); NIC59/JU1837 F1 female x JU1837 male cross divided by the percent viability of (JU1837); NIC59/JU1837 F1 female x JU1837 male cross). Cytoplasmic-nuclear combinations that show a statistically significant difference in viabilities between the two types of crosses and a relative viability <1 were considered to be cytoplasmic-nuclear

incompatibilities. Three biological replicates were performed for each cytoplasmic-nuclear combination except for JU1825 cytoplasmic - NIC24 nuclear and JU1825 cytoplasmic - NIC54 nuclear, which have four replicates each. For each biological replicate, 10 F1 hybrid L4 females were crossed to 10 L4 males on the same plate overnight at 25°C. The next day, they were moved to a new plate and allowed to lay embryos at 25°C for 1 hour. The parents were then removed and the percent viable progeny and embryonic lethality were calculated as described in the previous section of the Materials and Methods. The heat map used to visualize the median relative viability for each cytoplasmic nuclear combination was made using the heatmap.2 function from the gplot package in R.

2.4.4 *Molecular Methods*

To determine if either JU1825 or NIC59 are infected with *Wolbachia*, we performed PCR on crude lysates of both strains using degenerate primers targeted against two genes that are conserved in *Wolbachia* (Baldo *et al.* 2006). Specifically, we attempted to detect *gatB* (gatB_F1 with M13 adapter, TGTAAAACGACGGCCAGTGAKTTAAAYCGYGCAGGBGTT, and gatB_R1 with M13 adapter, CAGGAAACAGCTATGACCTGGYAAYTCRGGYAAAGATGA) and *fbpA* (fbpA_F3, GTTAACCCTGATGCYYAYGAYCC, and fbpA_R3, TCTACTTCCTTYGAYTCDCCRCC). As controls, we performed PCR on squash preps of *Drosophila melanogaster* w^{1118} mutant strains (Bloomington stock number 3605) that were infected or not infected with *Wolbachia*. *Drosophila melanogaster* strains were kindly provided by the laboratories of Harmit Malik and Leo Pallanck.

2.4.5 *Tetracycline treatment of JU1825 and NIC59*

Both JU1825 and NIC59 were passaged on 50 ug/mL tetracycline NGM plates streaked with OP50 for nine generations. Both strains were treated by crossing 10 L4 females and 10 L4 males on a fresh tetracycline plate each generation. Tetracycline plates were made by allowing NGM plates with OP50 lawns to soak up a mixture of tetracycline and 1x M9. The plates were left uncovered at room temperature until dry, and then used the following day.

2.4.6 *Statistics*

P values were determined using R (v 3.2.5). Several statistical tests were used (Kruskal-Wallis test followed by Dunn's test, and Fisher's exact test). When we performed several comparisons on the same dataset, we used the Bonferroni method to correct p-values for multiple testing. Most plots were made using the ggplot2 package in R.

2.5 ACKNOWLEDGEMENTS

We thank Marie-Anne Félix and Christian Braendle for providing the *Caenorhabditis nouraguensis* strains used in this study. We also thank the labs of Harmit Malik and Leo Pallanck for providing *Drosophila melanogaster* strains with and without *Wolbachia*. We thank Janet Young, Harmit Malik, Maitreya Dunham and Irini Topalidou for helpful discussions and comments on the manuscript. P.L. was supported in part by an NIH Institutional Training Grant (PHS, NRSA, T32GM007270 from NIGMS). This work was supported by an NSF CAREER Award (MCB-1552101) to M.A.

2.6 FIGURES

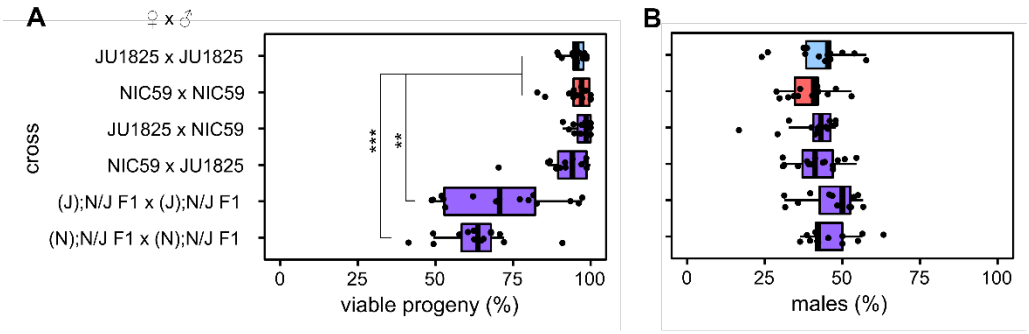


Figure 2.1. JU1825 and NIC59 exhibit F2 hybrid breakdown.

Crosses are listed on the y-axis. Letters in parentheses to the left of a semi-colon denote the cytoplasmic genotype of an individual (for example, "(J)" individuals have a JU1825 cytoplasmic genotype), while letters to the right of a semi-colon denote the genotypes of all autosomal loci (that is, "N/J" individuals are heterozygous NIC59/JU1825 throughout the autosomes). **(A)** Only (J); N/J F1 x (J); N/J F1 and (N); N/J F1 x (N); N/J F1 crosses exhibit a significant decrease in the percentage of viable progeny ($P<0.01$ and $P<0.001$, respectively). **(B)** There are no significant differences in the percentages of viable males between crosses ($P>0.05$). N=14 or 15 plates per cross. All p-values were calculated by a Kruskal-Wallis test followed by Dunn's test.

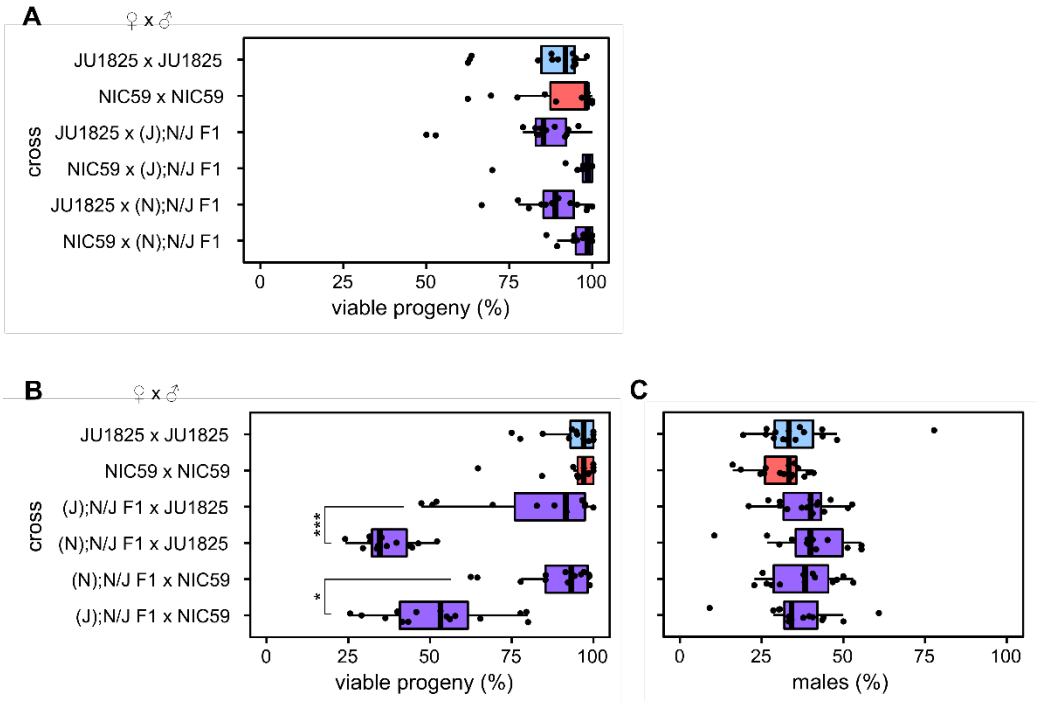


Figure 2.2. F2 inviability involves a maternal cytoplasmic effect.

(A) There is no significant difference in the percentage of viable progeny between any of the F1 hybrid male backcrosses and intra-strain crosses ($P>0.05$). **(B)** Backcrossing hybrid females to parental strain males reveals that only (N); N/J F1 female x JU1825 male crosses and (J); N/J F1 female x NIC59 male crosses exhibit a significant decrease in the percentage of viable progeny in comparison to intra-strain crosses ($P<0.001$). (N); N/J F1 female x JU1825 male crosses have significantly decreased viability in comparison to (J); N/J F1 female x JU1825 male crosses ($P<0.001$). Additionally, (J); N/J F1 female x NIC59 male crosses consistently have significantly decreased viability in comparison to (N); N/J F1 female x NIC59 male crosses ($P<0.05$). The viability of (J); N/J F1 female x JU1825 males can differ significantly between experiments (one of three biological replicates is shown here, see Figure 2.7 for the other two). **(C)** There are no significant differences in the proportion of viable males between the crosses ($P>0.05$). N=14 or

15 plates per cross. All p-values were calculated by a Kruskal-Wallis test followed by Dunn's test.

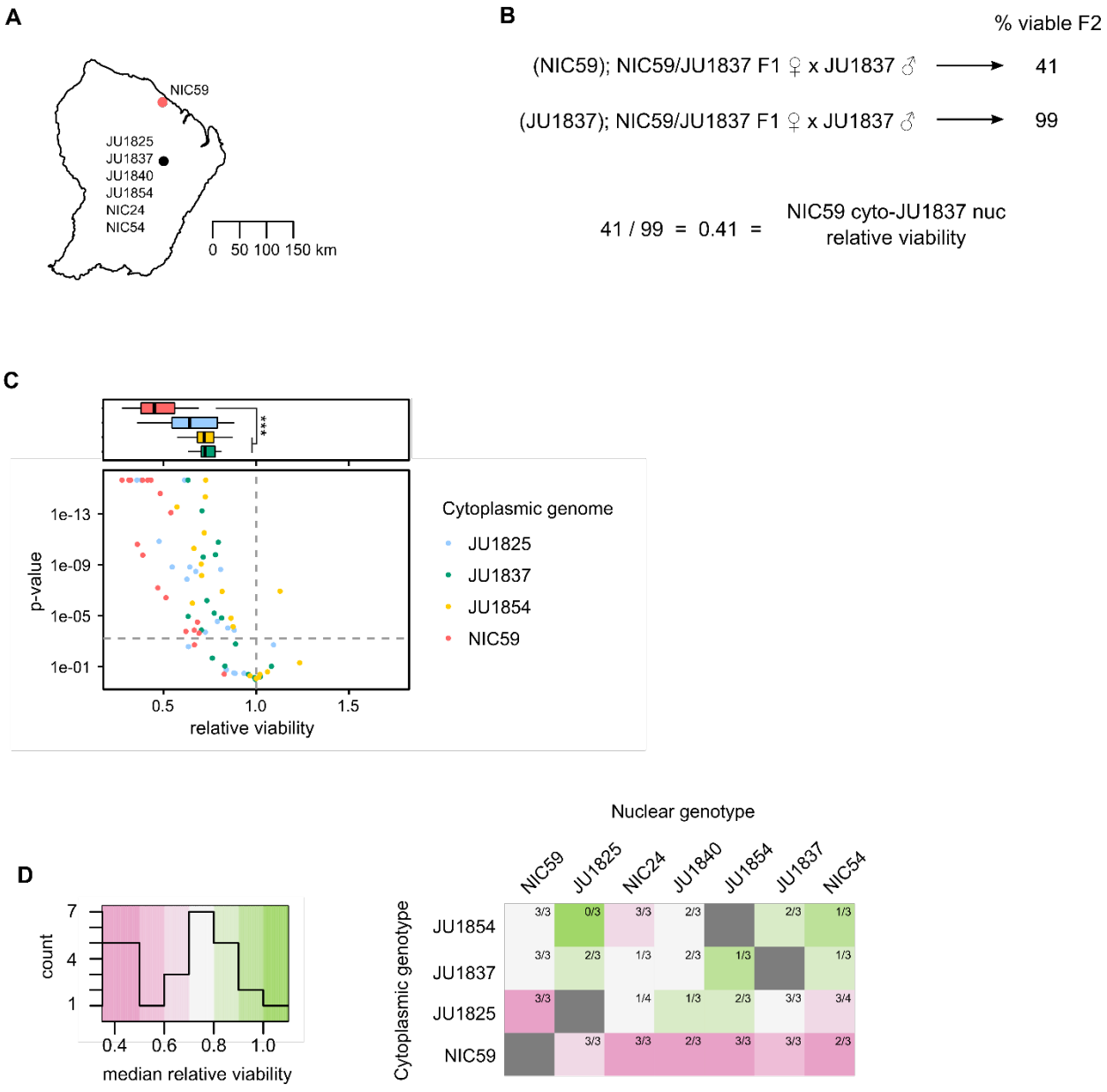


Figure 2.3. Cytoplasmic-nuclear incompatibility is widespread within *C. nouraguensis*. **(A)** A map depicting the two major sites where the strains used in this study were collected in French Guiana. GPS coordinates for NIC54 were obtained from Christian Braendle (personal communication), while the other six were obtained from Kiontke *et al.* 2011 and Félix *et al.* 2013. Strains in the southern collection site were collected from distinct rotten fruits or flowers within 2 km of each other and are represented as a single point. **(B)** To determine whether a

particular cytoplasmic-nuclear combination is incompatible, we tested for statistical differences in viability between the F1 female backcross that combines heterotypic cytoplasmic and nuclear genotypes (top cross) and the backcross that combines homotypic cytoplasmic and nuclear genotypes (bottom cross, see Materials and Methods). We also calculated the relative viability of the first cross to the second. **(C)** A scatter plot depicting all the cytoplasmic-nuclear compatibility tests performed. Each point corresponds to a single replicate of a certain cytoplasmic-nuclear combination. Points above the horizontal dashed gray line indicate statistically significant differences in viability between the two types of crosses mentioned in (B) ($P<0.0006$ after Bonferroni correction, Fisher's exact test). Points above the horizontal dashed gray line that have a relative viability <1 are considered statistically significant cytoplasmic-nuclear incompatibilities. The color of a point corresponds to the cytoplasmic genotype being tested. All cytoplasmic genotypes tested show an incompatibility with one or more heterotypic nuclear genotypes. See Figure 2.10 for separate graphs of all combinations. Above the scatterplot are boxplots depicting the relative viabilities of statistically significant cytoplasmic-nuclear incompatibilities. The color corresponds to cytoplasmic genotype tested. Incompatibilities involving the NIC59 cytoplasmic genotype have reduced viability compared to those involving the JU1837 and JU1854 cytoplasmic genotypes ($P<0.001$, Kruskal-Wallis test followed by Dunn's test). **(D)** A heatmap depicting the median relative viability for each cytoplasmic-nuclear combination. Each cytoplasmic-nuclear combination shows the proportion of replicates that exhibit significant incompatibilities (for example, 3 out of 3 replicates exhibit significant incompatibilities for the NIC59 cytoplasm–JU1854 nuclear combination, while only 1 out of 3 replicates exhibit significant incompatibilities for the JU1837 cytoplasm–JU1854 nuclear combination). Each cytoplasmic genotype is consistently incompatible with at least one

heterotypic nuclear genotype. The NIC59 cytoplasm has a distinct response to hybridization than the others tested.

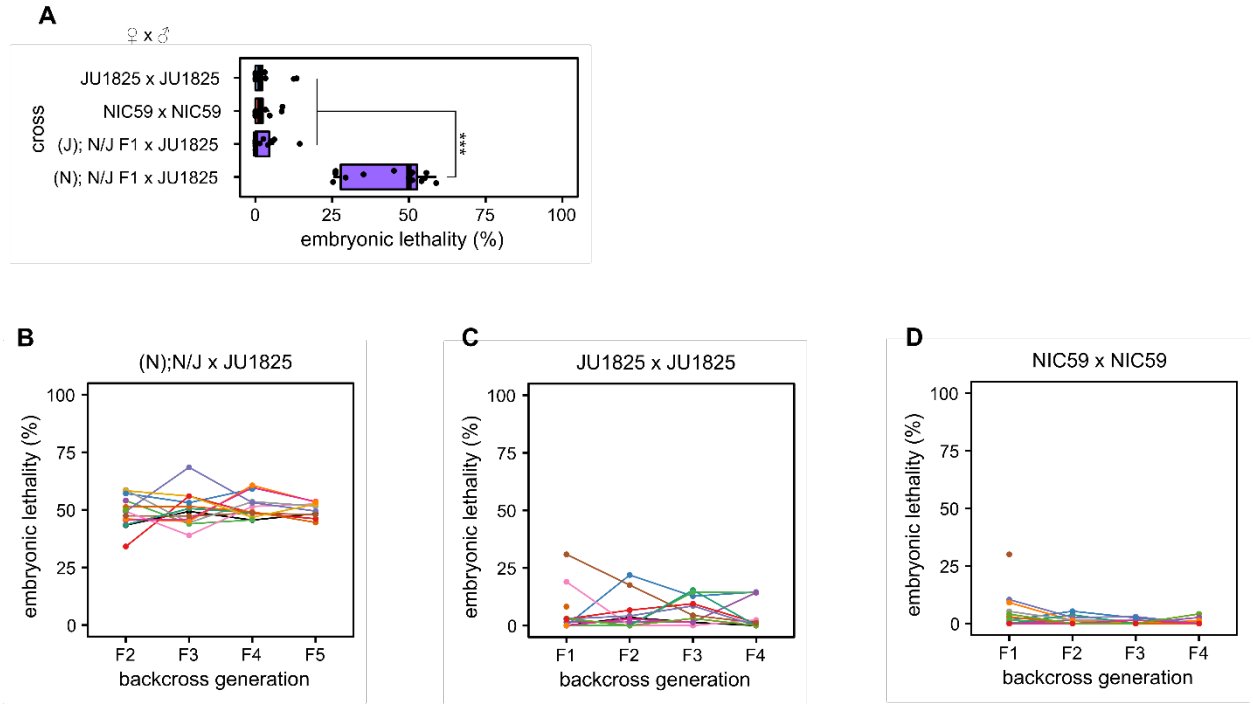


Figure 2.4. A single BDM incompatibility between a NIC59 cytoplasmic locus and a JU1825 nuclear locus causes embryonic lethality.

(A) Approximately 50% of the F2 progeny from (N); N/J F1 female x JU1825 male crosses arrest during embryogenesis, significantly higher than that seen in intra-strain crosses ($P<0.001$). In contrast, (J); N/J F1 female x JU1825 male and parental strain crosses exhibit similar low levels of embryonic lethality ($P>0.05$). N=14 or 15 plates per cross. **(B)** Initially, fifteen (N); N/J F1 females were independently backcrossed to single JU1825 males. For each independent lineage, a single surviving F2 female was again backcrossed to a JU1825 male. This backcrossing scheme was repeated until the F5 generation. Each colored line represents a single backcross lineage. All backcross lineages exhibit ~50% embryonic lethality throughout the backcross generations, consistent with the hypothesis that an incompatibility between a NIC59 cytoplasmic locus and a single JU1825 nuclear locus causes embryonic lethality. Number of independent backcross lineages assayed per generation: F2=15, F3=13, F4=13, F5=10. **(C)** The

JU1825 parental strain was “backcrossed” as a negative control. Number of independent backcross lineages assayed per generation: F1=15, F2=11, F3=11, F4=10. **(D)** The NIC59 parental strain was “backcrossed” as a negative control. Number of independent backcross lineages assayed per generation: F1=14, F2=12, F3=12, F4=12). All p-values were calculated by a Kruskal-Wallis test followed by Dunn’s test.

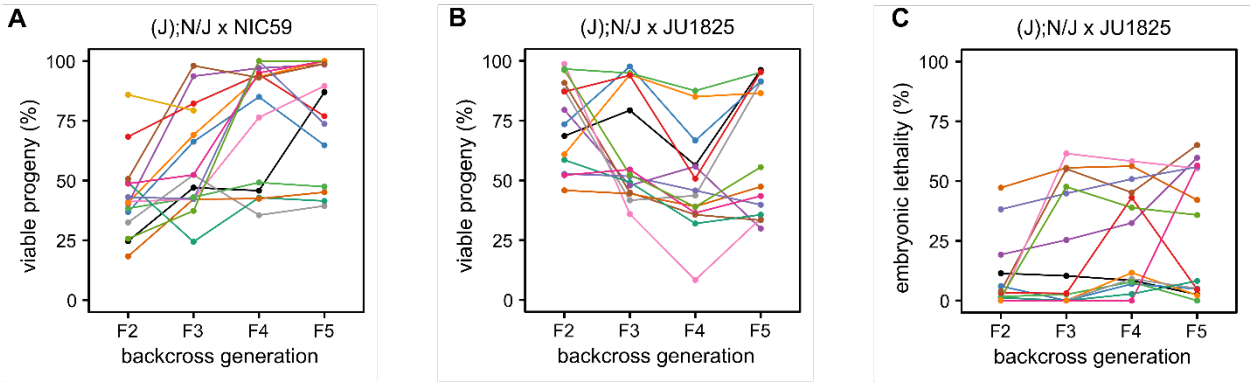


Figure 2.5. The JU1825 cytoplasm is heteroplasmic for JU1825-like and NIC59-like alleles. **(A)** The viability of independent (J); N/J female x NIC59 male backcross lineages were followed until the F5 generation. Surprisingly, in some lineages, multiple generations of backcrossing resulted in increased viability (similar to that seen in intra-strain crosses). Number of independent backcross lineages assayed per generation: F2=15, F3=15, F4=14, F5=14. **(B)** The viability of independent (J); N/J female x JU1825 male backcross lineages were also followed until the F5 generation. Interestingly, multiple generations of backcrossing resulted in some lineages with significantly reduced viability, similar to that seen in (N); N/J F1 female x JU1825 male crosses. Number of independent backcross lineages assayed per generation, F2 to F5=14. **(C)** Embryonic lethality of the same (J); N/J female x JU1825 male backcross lineages from Figure 2.5B (with same color-coding)**.** Upon additional generations of backcrossing, some (J); N/J female x JU1825 male lineages exhibit ~50% embryonic lethality, similar to (N); N/J F1 female x JU1825 male crosses. These results are consistent with the hypothesis that the JU1825 cytoplasm is heteroplasmic and contains JU1825-like and NIC59-like alleles.

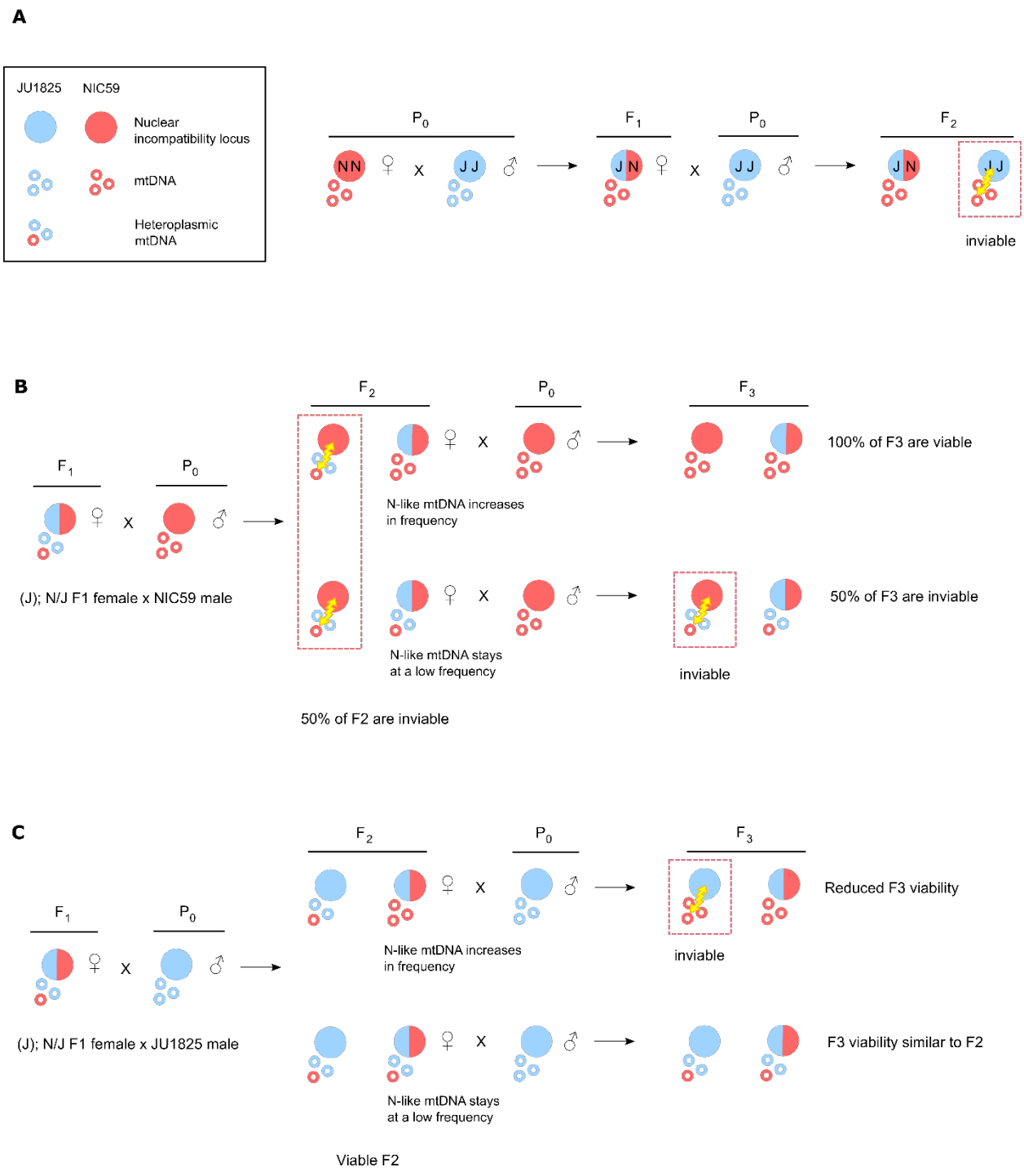


Figure 2.6. Mitochondrial-nuclear incompatibility model.

(A) We hypothesize that F2 hybrid breakdown is the result of a Bateson-Dobzhansky-Muller incompatibility between the NIC59 mitochondrial genome and a nuclear locus homozygous for

the JU1825 allele, and vice versa. As a specific example, when NIC59 females are crossed to JU1825 males, the resulting F1 hybrid females are expected to be heterozygous at all autosomal loci but inherit only NIC59 mtDNA. When F1 females are backcrossed to JU1825 males, F2 inviability results from an incompatibility between NIC59 mtDNA and an autosomal locus homozygous for the JU1825 nuclear allele. **(B)** We hypothesize that the JU1825 cytoplasm is heteroplasmic in F1 females and contains at least one NIC59-like allele. Backcrossing hybrid females with a JU1825 cytoplasm (that is, (J); N/J females) to NIC59 males for multiple generations can allow the NIC59-like cytoplasmic allele to increase in frequency and dilute out the effects of the incompatible JU1825 mtDNA (for example, top right F2 female). This eventually may allow the NIC59 nuclear locus to become homozygous and restore the viability of a lineage. On the other hand, the NIC59-like mtDNA can stay at a low frequency in viable F2 females (for example, bottom right F2 female). Backcrossing these F2 females to NIC59 males results in levels of inviability similar to the F2 generation. **(C)** By a similar line of reasoning, backcrossing hybrid females with a JU1825 cytoplasm to JU1825 males for multiple generations can allow the NIC59-like mtDNA to increase in frequency, where it can mimic the same genetic incompatibility seen in (N); N/J F1 female x JU1825 male crosses (Figure 2.6A).

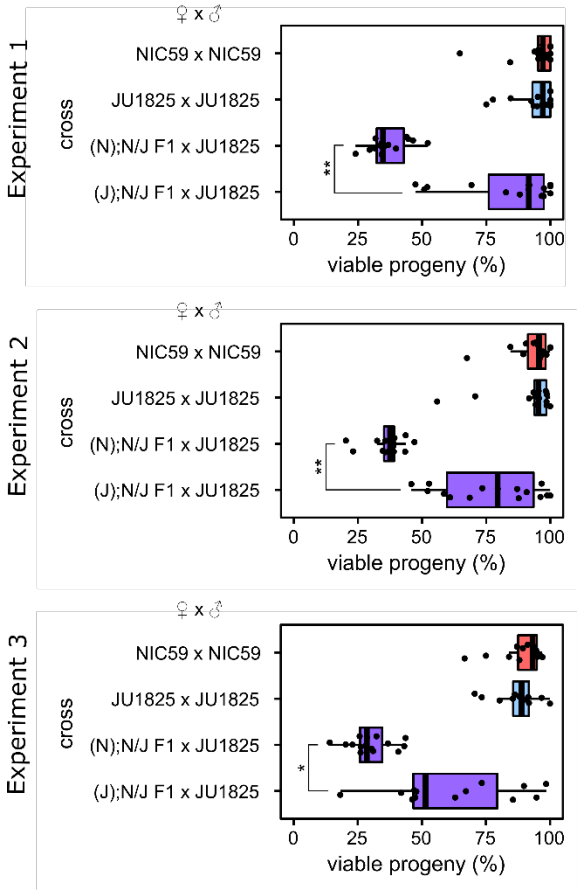


Figure 2.7. Variability of (J); N/J F1 female x JU1825 male crosses across experiments. Three biological replicates of the same type of backcross experiment. (J); N/J F1 female x JU1825 male crosses can either exhibit similar or significantly decreased rates of viability in comparison to intra-strain crosses (Experiment 1, non-significant, P>0.05; Experiment 2, non-significant, P>0.05; Experiment 3, P>0.05, non-significant in comparison to JU1825 x JU1825 crosses, P<0.05 significant in comparison to NIC59 x NIC59 crosses). However, (J); N/J F1 female x JU1825 male crosses consistently exhibit significantly increased rates of viability in comparison to (N); N/J F1 female x JU1825 male crosses (Experiment 1, **, P<0.01; Experiment 2, **, P<0.01; Experiment 3, *, P<0.05). Experiments 1 and 2 are data from Figures 2.2 and 2.5, respectively. All p-values were calculated by a Kruskal-Wallis test followed by Dunn's test.

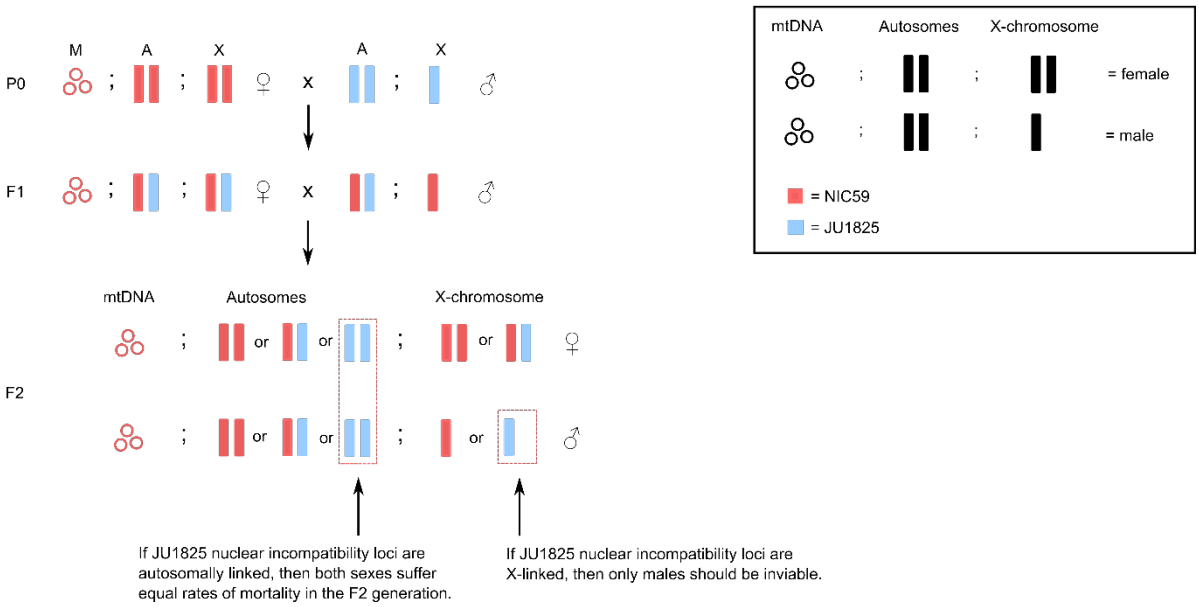
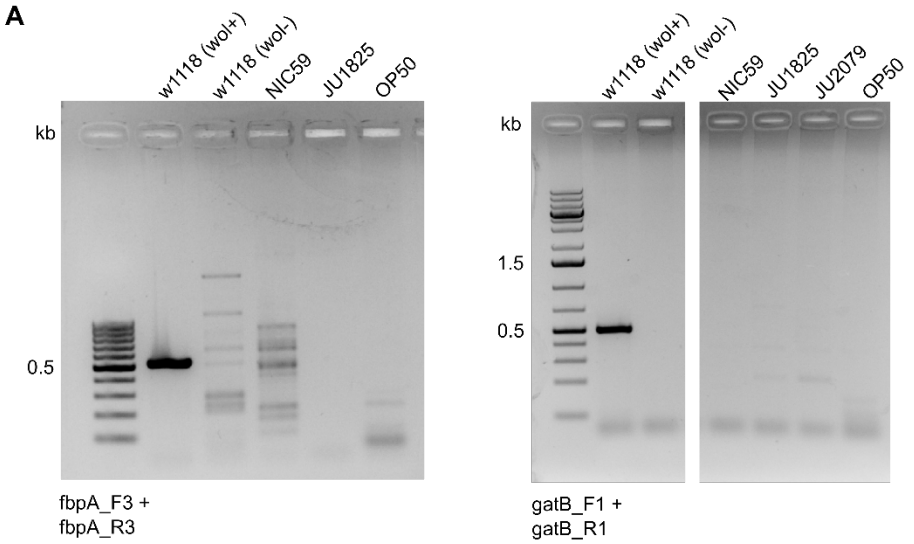


Figure 2.8. Nuclear incompatibility loci are linked to autosomes, not sex chromosomes.

F1 intercrosses allow us to infer that the nuclear incompatibility loci are autosomal, not X-linked. From the (N); N/J F1 x JU1825 male backcross experiment (Figure 2.2), we concluded that F2 inviability was the result of a genetic incompatibility between the NIC59 mitochondrial genome and nuclear loci homozygous or hemizygous for JU1825 alleles. It is reasonable to assume that the same genetic incompatibility contributes to F2 inviability in (N); N/J F1 female x (N); N/J F1 male crosses. If the nuclear incompatibility locus were X-linked, F2 male progeny of F1 intercrosses would have a 50% chance of being hemizygous for the JU1825 nuclear incompatibility locus whereas F2 females would only be heterozygous or homozygous for NIC59 alleles. Therefore, if the locus were X-linked, half of the F2 males would be inviable while females would be unaffected. If the nuclear incompatibility locus were autosomally linked, then both sexes would have an equal chance of being homozygous for the JU1825 nuclear incompatibility locus and thus, both sexes would be expected to suffer equal rates of inviability. We do not observe a significant decrease in the proportion of viable F2 males (Figure 2.1), so we

conclude that the JU1825 nuclear incompatibility locus or loci are linked to autosomes. The same line of reasoning can be used to show that the NIC59 incompatibility locus or loci are also autosomally linked.



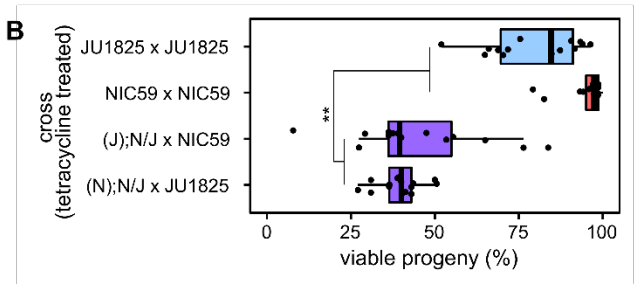


Figure 2.9. Endosymbiotic bacteria do not cause cytoplasmic-nuclear incompatibility. **(A)** PCR on both JU1825 and NIC59 crude lysates (10 adult worms per lysate, 5 females and 5 males) with degenerate primers against the *Wolbachia fbpA* or *gatB* loci fails to amplify the expected products. w^{1118} (wol+) and w^{1118} (wol-) *D. melanogaster* flies serve as positive and negative controls, respectively. PCR on crude lysates of OP50 (bacterial food source of NIC59 and JU1825) also fails to amplify the expected products. PCR on JU2079, an inbred strain derived from JU1825, also fails to amplify the expected *gatB* product. **(B)** After tetracycline treatment, both (J); N/J F1 female x NIC59 male and (N); N/J F1 female x JU1825 male crosses still exhibit significantly decreased levels of viability in comparison to tetracycline-treated intra-strain crosses ($P<0.01$). Additionally, there are no statistical differences in viability between

NIC59 x NIC59 and JU1825 x JU1825 tetracycline treated intra-strain crosses ($P>0.05$). N=14 or 15 for each cross. All p-values were calculated by a Kruskal-Wallis test followed by Dunn's test.

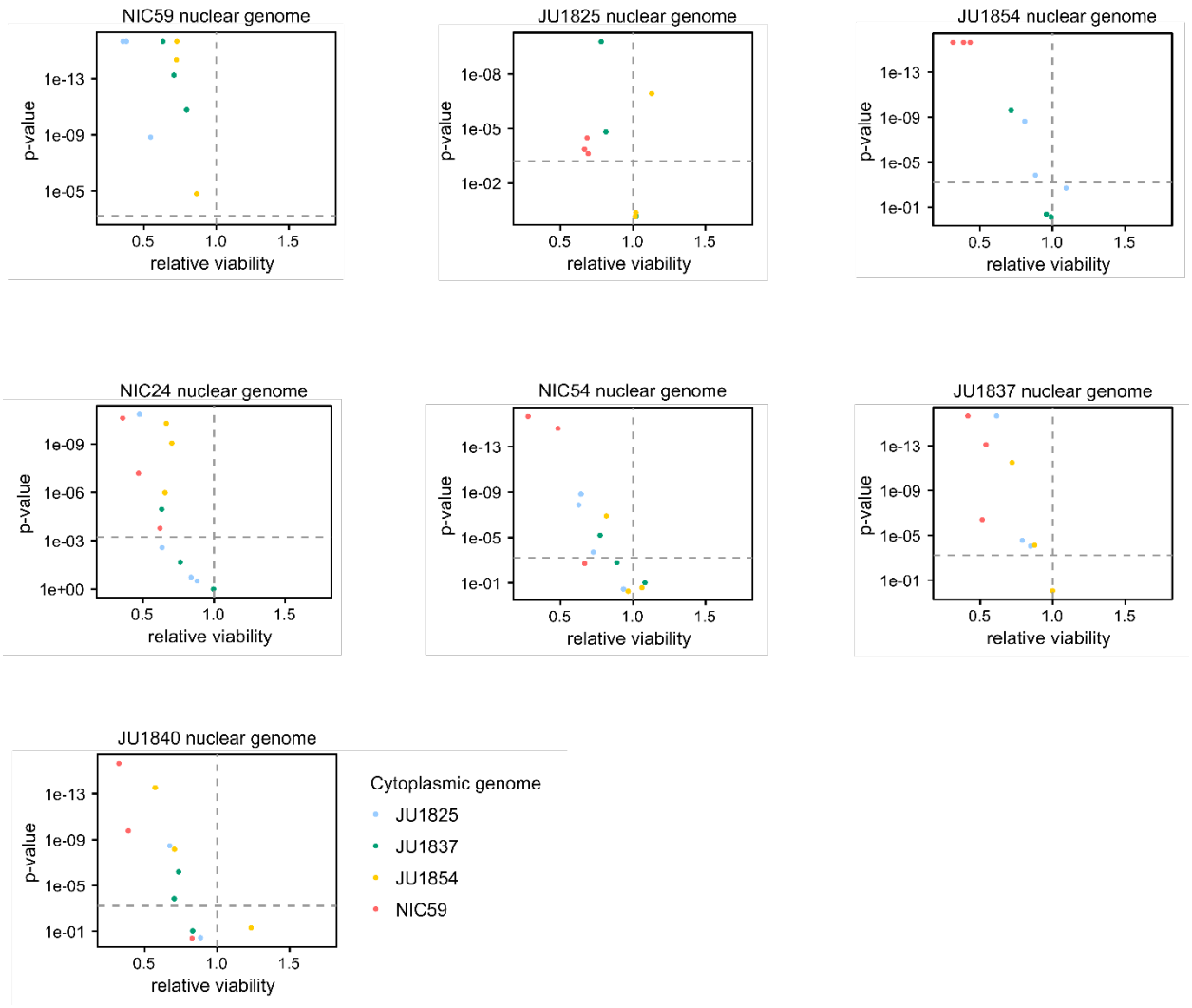


Figure 2.10. Cytoplasmic-nuclear tests of different *C. nouraguensis* strains.

Each graph depicts all the cytoplasmic-nuclear tests performed between four cytoplasmic genotypes and a single nuclear genotype. This is the same data that is grouped into a single graph in Figure 2.3C. Each cytoplasmic-nuclear combination has three biological replicates (except for JU1825 cytoplasm–NIC24 nuclear and JU1825 cytoplasm–NIC54 nuclear combinations, which have four replicates). Although there appear to be many cases of significant cytoplasmic-nuclear incompatibility (relative viability<1 and P<0.0006 after Bonferroni correction), there can be

discrepancies between replicates (for example, one replicate of the JU1825 cytoplasm–NIC24 nuclear combination indicates a significant incompatibility, while the other three do not).

Chapter 3. CYTOPLASMIC-NUCLEAR INCOMPATIBILITY BETWEEN WILD-ISOLATES OF CAENORHABDITIS NOURAGUENSIS – MAPPING THE CYTOPLASMIC AND NUCLEAR INCOMPATIBILITY LOCI

3.1 INTRODUCTION

We have shown that hybridizing two wild-isolates of *C. nouraguensis*, NIC59 and JU1825, results in a fraction of their F2 offspring dying before adulthood. We hypothesize that two cytoplasmic-nuclear incompatibilities strongly contribute to F2 inviability. The first incompatibility occurs between a JU1825 cytoplasmic locus (e.g. mtDNA) and an autosomal locus homozygous for NIC59 alleles, and kills half the F2 progeny derived from crossing (JU1825); NIC59/JU1825 F1 females to NIC59 males (Figure 3.1A). The second incompatibility occurs between a NIC59 cytoplasmic locus and an autosomal locus homozygous for JU1825 alleles, and kills half the F2 progeny derived from crossing (NIC59); NIC59/JU1825 F1 females to JU1825 males (Figure 3.1B). If our hypotheses are correct, then we predict that all viable F2 from these two crosses should be heterozygous (NIC59/JU1825) at the chromosome containing the nuclear incompatibility locus.

3.2 RESULTS

3.2.1 *The two nuclear incompatibility loci map to distinct chromosomes*

To test our hypothesis of how the genetic incompatibilities work, we pooled and whole genome sequenced a large population of viable adult F2 progeny from both backcrosses and

scanned their genomes for heterozygous regions (i.e. 0.5 NIC59 and JU1825 allele frequencies) (Figure 3.1A and Figure 3.1B). The *C. nouraguensis* reference genome assembly was generated from an inbred strain of JU1825, JU2079. Scaffolds of the JU2079 assembly were assigned to and ordered on chromosomes based on synteny to its sister-species, *C. becei* (see Chapter 4 Materials and Methods). Bulk sequencing of viable F2 from (JU1825); NIC59/JU1825 female x NIC59 male crosses show that only chromosome IV contains heterozygous regions, whereas all other chromosomes exhibit allele frequencies roughly expected of unlinked loci (i.e. ~0.75 NIC59 allele frequency, except chromosome V (0.69)) (Figure 3.1C). This result suggests that dead F2 are homozygous for NIC59 alleles at a locus on chromosome IV. Given this, we hypothesize that the JU1825 cytoplasm is incompatible with a homozygous NIC59 locus on chromosome IV and causes hybrid death. Bulk sequencing of viable F2 from (NIC59); NIC59/JU1825 female x JU1825 male crosses show that only chromosome III has regions that are heterozygous (Figure 3.1D). This result suggests that dead F2 are homozygous for JU1825 alleles at a locus on chromosome III. Given this, we hypothesize that the NIC59 cytoplasm is incompatible with a homozygous JU1825 locus on chromosome III and causes hybrid death. Taken together, the F2 bulk sequencing data indicate the two cytoplasmic-nuclear incompatibilities have genetically distinct nuclear components.

In the (N); N/J F1 female x JU1825 male backcross, chromosome IV, V and the X exhibit F2 allele frequencies consistent with Mendelian inheritance (0.25 N-allele frequency for the autosomes and 0.33 N-allele frequency for the X-chromosome). However, chromosome I and II exhibit an ~0.4 N-allele frequency, suggesting selection against JU1825 alleles. We hypothesize that the selection against JU1825 alleles on chromosome I and II result from a genetic incompatibility between NIC59 and JU1825 that results in the larval inviability seen in the (N);

N/J F1 female x JU1825 male cross (see Chapter 2). We hypothesize that the incompatibility causing larval inviability is genetically independent of the one that presumably causes F2 embryonic lethality (NIC59 cytoplasmic – JU1825 chromosome III). Furthermore, the genetics of the larval inviability incompatibility appear to be more complex than the embryonic lethal incompatibility, which seems to only require a single locus on chromosome III be homozygous for JU1825 alleles. The larval incompatibility might require a complex combination of chromosome I and II genotypes, however that combination is not immediately apparent from the bulk sequencing data.

3.2.2 *The JU1825 nuclear incompatibility locus is mapped to a 100 kb region on chromosome III*

F2 bulk sequencing mapped the nuclear incompatibility loci to single chromosomes. However, Mb-size regions on these chromosomes are heterozygous. To more finely map the JU1825 nuclear incompatibility locus, we introgressed the NIC59 cytoplasm into a JU1825 nuclear background. Specifically, we picked single (N); N/J hybrid females each generation for 15 generations and backcrossed them to JU1825 males (Figure 3.2A). Each backcross generation inherits the NIC59 cytoplasm and therefore viable offspring must be heterozygous at the nuclear incompatibility locus. However, there will be no selection maintaining heterozygosity at unlinked nuclear loci, which will eventually become homozygous for JU1825 alleles after multiple generations of backcrossing. Furthermore, recombination will break down the NIC59 chromosome III haplotype each generation and better resolve the nuclear incompatibility locus. At the F15 generation, introgression lines should be heterozygous NIC59/JU1825 only in regions linked to the JU1825 nuclear incompatibility locus. After the F15 generation, introgression lines underwent at least five generations of sib-mating to homozygose the nuclear incompatibility

locus for NIC59 alleles (however, see Materials and Methods). We then whole genome sequenced each introgression line and scanned for genomic regions that had a 1.0 NIC59 allele frequency.

Our most well-resolved introgression line (XZ1703) has a single 100 kb region homozygous for NIC59 alleles. The 100 kb region lies on a chromosome III scaffold, JU2079 scaffold19 (Figure 3.2B). This 100 kb NIC59 region was likely introgressed due to selection against JU1825 alleles at this locus, because XZ1703 still exhibits a cytoplasm that is incompatible with a single JU1825 nuclear locus (Figure 3.3C, cross (XZ1703); XZ1703/JU1825 female x JU1825 male). Therefore, this 100 kb region on JU2079 scaffold19 likely harbors the JU1825 nuclear incompatibility locus.

Allele frequencies in the 100 kb candidate locus in all but one introgression line are consistent with this region harboring the JU1825 nuclear incompatibility locus (see Materials and Methods). The one exception is introgression line XZ1705, which has an ~0% NIC59 allele frequency in the candidate region, but still had other regions on chromosome III that were homozygous for N-alleles (Figure 3.2B). We tested if XZ1705 has a cytoplasm that is incompatible with a nuclear locus homozygous for JU1825 alleles by comparing the viability of F2 offspring from these two crosses: (XZ1705); XZ1705/JU1825 F1 female x JU1825 male and (JU1825); XZ1705/JU1825 F1 female x JU1825 male (Figure 3.3C). The first cross resulted in 50% inviable F2 offspring whereas the second cross resulted in almost all F2 progeny surviving, therefore the XZ1705 cytoplasm is still incompatible with a nuclear locus homozygous for JU1825 alleles. However, the inviable progeny that result from this XZ1705 cytoplasmic – JU1825 nuclear incompatibility die primarily as larvae, which contrasts to the embryonic lethality seen in the NIC59 cytoplasmic – JU1825 nuclear and XZ1703 cytoplasmic – JU1825

nuclear incompatibilities. Interestingly, XZ1705 is the only introgression line with chromosome II loci that are homozygous for NIC59 alleles (data not shown) (all other introgression lines are homozygous for JU1825 alleles). As mentioned in the previous section, we hypothesized that F2 larval inviability in crosses between NIC59 and JU1825 result from a genetic incompatibility between NIC59 and JU1825 involving chromosomes I and II (Figure 3.1D). Perhaps the larval inviability observed in crosses between XZ1705 and JU1825 is caused by an incompatibility with a similar genetic basis. However, it is interesting XZ1705 is also homozygous for NIC59 alleles on chromosome III, suggesting selection against JU1825 alleles. Regardless, given that introgression line XZ1705 has a distinct lethal phenotype when hybridized to JU1825 and because it is the only introgression line that does not have NIC59 alleles in the candidate 100 kb region, we hypothesized that it exhibits a cytoplasmic-nuclear incompatibility that is genetically distinct from the NIC59 cytoplasmic – JU1825 nuclear incompatibility that causes embryonic lethality.

To test if the XZ1705 cytoplasmic – JU1825 nuclear and XZ1703 cytoplasmic – JU1825 nuclear incompatibilities are genetically distinct, we determined whether XZ1705 and XZ1703 contain NIC59 nuclear loci that can rescue the F2 inviability of both incompatibilities (Figure 3.3A and 3.3B). For example, backcrossing (XZ1703); XZ1703/JU1825 F1 females to JU1825 males results in 50% F2 embryonic lethality, with F2 individuals homozygous for JU1825 alleles at the nuclear incompatibility locus dying during embryogenesis and heterozygous individuals (N/J) surviving to adulthood. However, backcrossing (XZ1703); XZ1703/JU1825 females to XZ1703 males would result in F2 offspring either being heterozygous (N/J) or homozygous NIC59 at the nuclear incompatibility locus, and therefore rescue F2 inviability (i.e. ~0% embryonic lethality). This is what we observe (Figure 3.3). If XZ1705 is also homozygous for

NIC59 alleles at the nuclear incompatibility locus that causes embryonic lethality, then backcrossing (XZ1703); XZ1703/JU1825 F1 females to XZ1705 males will also result in 0% F2 embryonic lethality. If XZ1705 does not have NIC59 alleles at the nuclear incompatibility locus that causes embryonic lethality, then F2 embryonic lethality will not be rescued (i.e. 50% F2 embryonic lethality). We find that XZ1705 does not rescue F2 embryonic lethality (i.e. (XZ1703); XZ1703/JU1825 F1 x XZ1705 male results in 50% F2 embryonic lethality). We then performed an analogous cross to show that XZ1703 fails to rescue the larval lethality of the XZ1705 cyto – JU1825 nuclear incompatibility (i.e. (XZ1705); XZ1705/JU1825 F1 female x XZ1703 male results in ~50% larval lethality). Therefore, the data suggest that the XZ1705 cytoplasmic – JU1825 nuclear incompatibility is genetically distinct from the XZ1703 cytoplasmic – JU1825 nuclear incompatibilities. Therefore, the 100 kb candidate region mapped in XZ1703 likely harbors the JU1825 incompatibility locus required for the original embryonic lethal NIC59 cytoplasmic – JU1825 nuclear incompatibility.

Introgression line XZ1705 is also strange in that even though it should have NIC59 mitochondrial DNA, it is homozygous for JU1825 alleles at our hypothesized embryonic lethal incompatibility locus. If there indeed is an incompatibility between the NIC59 mitochondrial genome and homozygosity of JU1825 alleles at the 100 kb candidate locus, then presumably XZ1705 has suppressed the incompatibility in some way. Suppression could potentially occur through changes at the NIC59 cytoplasmic locus (e.g. the NIC59 mtDNA) or the JU1825 nuclear locus. Comparing these sequences in XZ1705 to other introgression lines could potentially give insight into what loci are required for embryonic lethality.

3.2.3 *100 kb candidate region contains tandem repeats of F-box genes, transmembrane domain containing genes, leu-tRNAs and nuclear hormone receptors.*

The JU1825 nuclear incompatibility locus was mapped to JU2079 scaffold19. Zooming in on JU2079 scaffold19, we find that the candidate region is approximately 100 kb in length, with introgression line XZ1784 defining the right end and XZ1703 defining the left end (Figure 3.4A).

JU2079 scaffold19 contains several clusters of tandem duplications (Figure 3.4B). The first cluster (18 kb – 90 kb) consists of an approximately 1.5 kb repeating unit, each of which contains roughly one F-box domain containing gene. F-box genes are known to be adapters for ubiquitin mediated proteolysis, with their F-box domain interacting with subunits of an CUL-1 E3-complex and their C-termini interacting with substrates that are targeted for ubiquitination and degradation (Skaar *et al.* 2013). This cluster encodes 40 F-box genes, with the closest paralogs within the cluster being ~97% identical in amino acid sequence. Aligning the amino acid sequences of these paralogs shows they have a relatively conserved F-box domain, but their N-termini and pockets of their C-termini can be quite variable (Figure 3.5B). The second cluster (107 kb – 116 kb) consists of an approximately 1.6 kb repeating unit. Within this cluster are 13 transmembrane domain containing genes (4-6 transmembrane domains each), with the closest paralogs in this cluster being ~75% identical in amino acid sequence. These transmembrane containing genes do not have any other known motifs and do not have any obvious homologs outside of *C. nouraguensis*' sister-species (*C. becei*). Within this cluster are also three leucine tRNAs that have identical DNA sequences. These first two clusters are roughly within the 100 kb candidate region. There are also three nuclear hormone receptor genes in the candidate region (Figure 3.5A). The third major repetitive cluster (138 kb – 164 kb), which is outside of the candidate region, consists of a 1.6 kb repeating unit which also contains many F-box genes.

Because an individual with a NIC59 cytoplasmic genotype needs at least one NIC59 allele present at the 100 kb candidate locus to be viable, then NIC59 encodes something within the candidate region required for viability whereas JU1825 does not. To determine what this NIC59 specific locus could be, we compared the NIC59 and JU1825 genome assemblies in the 100 kb candidate region. Specifically, we aligned scaffolds from both the NIC59 and JU1825 assemblies to the candidate region in the JU2079 assembly using MUMMER (Kurtz *et al.* 2004). NIC59 and JU1825 are outbred strains and contain lots of heterozygosity and therefore their assemblies are fragmented into smaller scaffolds than the inbred JU2079 assembly. Many small scaffolds from the JU1825 genome align to the JU2079 candidate region, making it difficult to determine all genes encoded in this region (Figure 3.4E). Luckily, we determined that the NIC59 cytoplasm is also incompatible with JU2079 chromosome III (Figure 3.4C), and therefore genes in the JU2079 candidate region can be compared to the homologous NIC59 sequence. Luckily, the NIC59 assembly is not as fragmented, and several relatively large NIC59 scaffolds map to the candidate region (Figure 3.4D, Table 3.2). However, some NIC59 scaffolds are likely misaligned because only a small fraction of their sequence map to the candidate region (NIC59 scaffold625: 0.12% of 97 kb, scaffold1330: 1.6% of 50 kb, scaffold730: 2.65% of 33 kb, scaffold2903: 0.36% of 31 kb). In contrast, other NIC59 scaffolds have large portions which map to the JU2079 candidate region (e.g. NIC59 scaffold4749: 61% of 12 kb) and are likely from the same genomic region in NIC59.

We find that the NIC59 scaffolds that map to the candidate region also encode F-box, nuclear hormone receptor, transmembrane domain and leu-tRNA genes (Figure 3.5B). We find slightly less F-box and nuclear hormone receptor genes in NIC59 compared to JU2079 (i.e. NIC59 has 29 F-box genes whereas JU2079 has 40 and NIC59 has six nuclear hormone receptor

genes whereas JU2079 has eight). To aid in our search for NIC59 specific loci that are required for viability, we made a maximum-likelihood phylogeny of all whole F-box and nuclear hormone genes encoded by both the NIC59 and JU2079 assembly, and looked for NIC59 genes that were particularly divergent from their JU2079 homolog or do not have an immediate JU2079 homolog. We found that each NIC59 nuclear hormone receptor gene has a clear homolog in JU2079 (data not shown). However, we found six NIC59 F-box genes do not have an immediate JU2079 homolog, and therefore are potential candidates (Figure 3.6).

Many small NIC59 scaffolds map to repeat cluster containing the transmembrane domain genes, inhibiting our ability to determine the amino acid sequence of these NIC59 transmembrane genes (Figure 3.4D, Figure 3.7B and Table 3.2). Interestingly, mapping NIC59 and JU1825 reads to the JU2079 assembly shows that NIC59 has much higher read depth and therefore many more copies of a subset of transmembrane domain and tRNA genes in comparison to JU1825 and JU2079 (Figure 3.5C). Based on read depth, NIC59 encodes ~73 transmembrane domain genes and ~28 leu-tRNAs, whereas JU2079 encodes only 13 transmembrane domain genes and 3 leu-tRNAs (Figure 3.5B).

3.2.4 *Making and screening the NIC59 fosmid library*

To further narrow down the NIC59 locus required for viability, we planned to make NIC59 fosmid clones that span the candidate region and test whether they rescue the inviability of F2 hybrids that have a NIC59 cytoplasmic genotype and are homozygous for JU1825 alleles at the candidate locus. Each fosmid should be between 30-40 kb in size, therefore it would take at least four clones to span the entire candidate region. Fosmid clones that rescue F2 inviability would then be digested into smaller segments and subcloned. These subclones would then be

tested for their ability to rescue F2 inviability. Subcloning would be done until a small and manageable number of genes reside in a rescuing subclone.

We generated a fosmid library of the entire NIC59 genome (Figure 3.7A and Materials and Methods). The library is composed of three deep-well 96-well plates, with each well containing a pool of ~50 NIC59 fosmid clones frozen as a bacterial glycerol stock. Therefore, any NIC59 genomic locus should on average have at least 5.7-7.6 fosmid clones covering it.

We PCR screened each well of the fosmid library with primers that amplify sequences on NIC59 scaffolds that span the candidate region (Figure 3.7B). Because the region is so repetitive, we made sure that these primers annealed to unique sequences, and as a control showed that each set of primers amplified a PCR product when using NIC59 genomic DNA as a template (data not shown). At least one set of screening primers were made for each NIC59 scaffold. Of the 11 sets of screening primers, only four amplified their target/template DNA from the NIC59 fosmid library (Figure 3.7B). Furthermore, each of these four sets of primers only amplified a PCR product from a single well, indicating that there is only one fosmid mapping that locus, rather than the expected 5-7 (oPL340+341 and oPL342+343: Plate1Well9H, oPL346+347: Plate2Well3C, oPL348+349: Plate2Well2B).

We then tried to isolate single PCR positive clones from each well. First, we plated and PCR screened single colonies from Plate1Well9H using primers oPL340+341 and oPL342+343. Based on our titer when initially making the library, we believed that there were roughly 50 different fosmid clones per well, and therefore expected that ~1/50 colonies would give a PCR positive signal. However, we found that our fosmid of interest is at much lower frequency. We initially found this by PCR genotyping 300 colonies from Plate1Well9H and finding none were positive. To get a sense of the clone's actual frequency, we made subpools of Plate1Well9H. We

grew up a 4ml overnight of Plate1Well9H, calculated the overnight's titer, and made 96 subpools in a 96-well plate, with each well/subpool having 200 clones (Figure 3.7C). We grew these subpools overnight and PCR screened them the next day. Only three subpools were PCR positive, indicating the fosmid of interest is at a frequency of 1/6,000 in the overnight of the Plate1Well9H glycerol stock (Figure 3.7C). Finding our fosmid at such a low frequency either suggests that the initial titer when making the library was incorrect and/or there is selection against clones in this region. To determine if our fosmid of interest is selected against, we made another set of subpools (round two subpools) from one of the PCR positive subpools and grew them up overnight. Each of the round two subpools had approximately 200 clones/well. We found that 24/96 round two subpools were PCR positive and therefore the frequency of the fosmid of interest is 1/800 (Figure 3.7C). Because the frequency of the fosmid fell from 1/200 to 1/800, we believe there is selection against our fosmid of interest when growing them overnight. We believe this selection could partially explain the low frequency of our fosmid in Plate1Well9H (i.e. 1/6,000). Specifically, there is one overnight growth step required for making the glycerol stocks of the fosmid library and another overnight growth step required for plating out bacteria for single colony PCR. Therefore, this fosmid could decrease in frequency from 1/50 when the initial library was made to 1/800 when plating bacteria out for single colony PCR.

We also tried isolating our fosmid of interest from Plate1Well9H using colony hybridization. We generated a PCR-DIG labelled oligo probe using oligos modified from oPL340 and oPL341, and probed a plate with ~300 colonies (see Materials and Methods). We found that the probe was very non-specific, with many colonies lighting up (Figure 3.7B). However, we picked and PCR screened a subset of colonies that seemed to give a relatively stronger signal. One of these colonies was positive for both oPL340+341 and oPL342+343

(Figure 3.7B). We isolated our clone of interest and end-sequenced its insert and found that one end mapped to one of the scaffolds in the 100 kb candidate region. Surprisingly, we found that the other end of the insert mapped to scaffold1893 of the NIC59 genome assembly, which is homologous to a JU2079 scaffold that does not map to chromosome III (i.e. JU2079 scaffold362, which maps to chromosome I). At first, we hypothesized that the NIC59 genome was rearranged relative to the JU2079 genome and that NIC59 scaffold1893 did lie within the candidate region. However, all introgression lines are homozygous for JU1825 alleles on JU2079 scaffold362 (data not shown), indicating that NIC59 scaffold1893 does not reside in the candidate region. Instead, we suspect these two sequences were artificially fused together while making the fosmid library.

3.2.5 *Mapping the NIC59 cytoplasmic incompatibility locus*

Because we hypothesized that the cytoplasmic incompatibility loci could be encoded by the mitochondrial genome, we set-out to find functional differences between the NIC59 and JU1825 mitochondrial genomes. The *C. nouraguensis* mitochondrial genome encodes 12 protein subunits of the electron transport chain, 22 tRNAs and 2 rRNAs (Figure 3.8A). The NIC59 and JU1825 mitochondrial genomes differ by 95 SNPs and a 1 bp indel. Potential functional differences between the two strains include single non-synonymous SNPs in nd-1 (L242F) and cox-1 (V426I) (subunits of complex I and IV of the electron transport chain, respectively), one SNP in the cysteine-tRNA gene and 3 SNPs and a 1 bp indel in the 16S rRNA gene (Figure 3.8A).

Despite finding potential functional differences between the NIC59 and JU1825 mitochondrial genomes, editing mtDNA currently not possible in *Caenorhabditis*. Therefore, we

can only use transgenic experiments to directly test whether protein coding mtDNA loci are the cytoplasmic component of the incompatibility. However, the unexpected results from the backcross experiments mentioned in Chapter 2 might aid in determining whether the mitochondrial genome in fact harbors the cytoplasmic incompatibility locus.

As previously mentioned, backcrossing (N); N/J F1 females to JU1825 males results in approximately 50% of their F2 progeny dying during embryogenesis, presumably due to a NIC59 cytoplasmic – JU1825 nuclear incompatibility. In contrast, backcrossing (J); N/J F1 females to JU1825 males results in almost no F2 offspring dying during embryogenesis, presumably because the JU1825 cytoplasm is compatible with its own nuclear loci. But unexpectedly, backcrossing post-F2 generation (J); N/J females to JU1825 males for additional generations produces a subset of lineages that suddenly start to exhibit ~50% embryonic lethality (Figure 3.8B). To determine whether embryonic lethality observed in these (J); N/J female backcrosses also results from the death of individuals homozygous for JU1825 alleles on chromosome III, we PCR genotyped each (J); N/J female that gave rise to each backcross generation at a locus on chromosome III. This experiment was performed before sequencing the introgression lines, so the genotyping marker lies on a scaffold in the center of chromosome III but not on the same scaffold as the candidate region. We find that backcross lineages that exhibit 50% embryonic lethality are almost always heterozygous for chromosome III, while those that show low levels can either be heterozygous or homozygous JU1825. By the F5 generation, we find that 6/7 backcross lineages that exhibit 50% embryonic lethality were heterozygous at our chromosome III marker, whereas 9/9 backcross lineages that exhibit 0-12% embryonic lethality were homozygous for JU1825 alleles at our chromosome III marker (Figure 3.8B). We hypothesize that the same genetic incompatibility seen between the NIC59 cytoplasm and

JU1825 nuclear genome that causes embryonic lethality is also occurring in the subset of (J); N/J female x JU1825 male backcross lineages that exhibit 50% embryonic lethality.

If the NIC59 mitochondrial genome harbors the cytoplasmic incompatibility locus, then we predict that the NIC59 mitochondrial genome should be present in the subset of (J); N/J female x JU1825 male backcross lineages that exhibit 50% embryonic lethality. However, (J); N/J females should theoretically only have inherited JU1825 mtDNA from their JU1825 mother. One possibility is that (J); N/J females inherited some of their NIC59 father's mtDNA (paternal leakage) and have a J/N heteroplasmic genotype. Then presumably the (J); N/J backcross lineages that exhibit high embryonic lethality are enriched for the NIC59 mtDNA whereas those that have low embryonic lethality have no or lower amounts. To test this hypothesis, we PCR genotyped all F4 females that gave rise to the F5 backcross generation at a mtDNA locus (oPL205+54, *cox-1, cys-tRNA, met-tRNA, asp-tRNA, gly-tRNA and cox-2*). The PCR assay should only amplify NIC59 alleles and therefore should detect small amounts of NIC59 mtDNA in a heteroplasmic situation. None of the females gave a positive signal and therefore do not carry NIC59 mtDNA (Figure 3.8C). Therefore, we can eliminate paternal leakage of NIC59 mtDNA as the cytoplasmic component of the incompatibility seen in (J); N/J backcross lineages.

3.3 DISCUSSION

In this study, we attempted to map the cytoplasmic (e.g. mtDNA) and nuclear components of a deleterious genetic incompatibility between two wild-isolates of *C. nouraguensis,* NIC59 and JU1825. By sequencing large populations of viable F2 progeny, we found that the JU1825 cytoplasm is incompatible with a locus or loci on NIC59 chromosome IV (NIC59 nuclear incompatibility locus), and that the NIC59 cytoplasm is incompatible with a

locus or loci on JU1825 chromosome III (JU1825 nuclear incompatibility locus). Because hybrid death does not occur unless the nuclear incompatibility locus is homozygous, we hypothesize that the nuclear incompatibility locus fails to provide a function required for viability. By sequencing introgression lines, we finely mapped the JU1825 nuclear incompatibility locus to an approximately 100-kb candidate region on JU2079 scaffold19. In the JU2079 assembly, the candidate region contains two major clusters of tandem duplications, one containing mostly F-box genes with some nuclear hormone receptors and the other containing mostly transmembrane genes and some leucine tRNAs. When comparing the candidate region between the NIC59 and JU2079 assemblies, we find that NIC59 has six F-box genes with no immediate JU2079 homologs and a dramatically increased number of transmembrane domain genes. To determine which NIC59 genes in the candidate region are required for viability, we attempted to generate NIC59 fosmid clones spanning the candidate region with the hope that some would rescue the inviability of F2 hybrids homozygous for JU1825 alleles at the incompatibility locus. However, technical difficulties in NIC59 fosmid library production and potential selection against fosmid clones have halted our efforts.

We have made progress in understanding the NIC59 cytoplasmic incompatibility locus. Specifically, we show that the NIC59 cytoplasmic – JU1825 nuclear incompatibility begins to emerge when (JU1825); NIC59/JU1825 females are backcrossed to JU1825 males for multiple generations. We believe this is the case because maintenance of heterozygosity at a chromosome III locus in females that found each backcross generation is associated with 50% embryonic lethality, suggesting individuals that are homozygous for JU1825 alleles at chromosome III are inviable. However, we fail to detect NIC59 mtDNA in any of the (JU1825); NIC59/JU1825

female x JU1825 male backcross lineages, suggesting that NIC59 mtDNA does not harbor the cytoplasmic incompatibility locus.

3.3.1 *Genetically distinct nuclear incompatibility loci.*

Despite NIC59 and JU1825 both having a cytoplasm that is incompatible with the other's nuclear genome, we find that the nuclear incompatibility loci map to different chromosomes and are likely genetically distinct. This is not particularly surprising, considering most molecularly characterized incompatibility loci act asymmetrically, meaning that only one of the two alleles (e.g. NIC59 or JU1825) at a locus are incompatible with heterospecific loci (Brideau *et al.* 2006; Lee *et al.* 2008; Ferree and Barbash 2009; Chou *et al.* 2010). This is also the case in known mitochondrial-nuclear incompatibilities. For example, the mitochondrial genomes of hermaphroditic plants often evolve novel ORFs that prevent the production of pollen and allocate more resources to the production of ovules (i.e. cytoplasmic male sterility or CMS) (Chase 2007). This benefits the mitochondrial genome because it is only inherited through ovules, but at odds with the nuclear genome which is equally inherited through both pollen and ovules. Therefore, nuclear suppressors of CMS that rescue male fertility will be selected for and increase in frequency. These nuclear suppressors are often pentatricopeptide containing proteins which prevent the translation of CMS inducing mitochondrial ORFs (Fujii *et al.* 2011). Isolated plant populations often undergo independent bouts of co-evolution between CMS inducing ORFs and nuclear suppressors that rescue male fertility. Therefore, when hybridizing two different species or populations of the same species, each nuclear genome will likely lack suppressors of heterospecific CMS inducing mtDNAs (Burt and Trivers 2006). Therefore, even though hybridization between these two populations produces two cytoplasmic-nuclear incompatibilities

(i.e. CMS), the nuclear loci are genetically distinct. We hypothesize that at least two distinct bouts of cytoplasmic-nuclear co-evolution have occurred in *C. nouraguensis* to result in the genetically distinct cytoplasmic-nuclear incompatibilities.

3.3.2 *The nuclear incompatibility locus*

Since being homozygous for JU1825 alleles at the nuclear incompatibility locus results in hybrid death, we hypothesize that the homologous NIC59 region contains loci required for viability. However, considering that the candidate region is large and contains many genes that are potentially functionally divergent between NIC59 and JU2079 (based on amino acid identity), it is currently unclear whether one or more NIC59 genes are required for viability. Indeed, incompatibility loci that cause hybrid sterility and inviability in other systems can require the presence of multiple neighboring genes (Long *et al.* 2008; Seidel *et al.* 2011; Liénard *et al.* 2016). Aside from functional divergence of amino acid sequences, the differences in copy number between NIC59 and JU2079 could potentially underlie the incompatibility. Specifically, NIC59 encodes many more transmembrane and leucine tRNA genes than JU2079 and it might be that their higher dosage is required for viability. Interestingly, one potential functional difference between the NIC59 and JU1825 mitochondrial genomes is a SNP in a cysteine tRNA. In marsupials it has been shown that a nuclear encoded lysyl tRNA is imported into the mitochondria to compensate for a pseudogenized mitochondrial lysyl tRNA (Dörner *et al.* 2001). Despite encoding a different tRNA, perhaps overexpression of the leucine tRNA in the candidate region and its import into the mitochondria compensates for a suboptimal NIC59 mtDNA cysteine tRNA.

One of the most striking features of the candidate region is that it bears a remarkable resemblance to the *zeel-1 peel-1* locus in *C. elegans* (Seidel *et al.* 2008, 2011). In *C. elegans*, the *zeel-1 peel-1* locus is known to cause incompatibilities when crossing wild-type Bristol (z+p+) and Hawaiian (Δzp, does not encode *zeel-1 peel-1*) strains. PEEL-1 is a small transmembrane containing gene with no known homologs outside of *C. elegans*. ZEEL-1 is closely related to a family of proteins that resemble a substrate recognition subunit of a CUL-2 E3 ubiquitin ligase complex. In F1 z+p+/ Δzp hermaphrodites, peel-1 is expressed in the male germline and packaged into sperm, which upon fertilization delivers PEEL-1 protein to F2 embryos. F2 progeny that fail to inherit the *zeel-1 peel-1* locus (Δzp/Δzp) die during embryogenesis. In contrast, F2 progeny that inherit the *zeel-1 peel-1* locus express *zeel-1* during embryogenesis, which rescues the lethal effects of PEEL-1. It seems that PEEL-1 acts as a poison and ZEEL-1 acts as its antidote, and ZEEL-1 might suppress PEEL-1 toxicity by promoting its degradation via ubiquitination. It is interesting that my candidate region also encodes adapters for ubiquitin mediated degradation (F-box genes) as well as transmembrane genes with no homology to anything outside of *C. nouraguensis*' sister-species (*C. becei*), potentially indicating that a toxin antidote system underlies the inviability seen in F2 hybrids between NIC59 and JU1825. If this were the case, we hypothesize a toxic NIC59 transmembrane gene would be expressed in the germline of NIC59/JU1825 F1 females and distributed to all F2 embryos. Then only F2 embryos inheriting NIC59 alleles in the candidate region can express a NIC59-specific F-box gene during embryogenesis and suppress the toxic effects of the NIC59 transmembrane gene. However, this model alone fails to explain the cytoplasmic effect of the incompatibility. To be consistent with the cytoplasmic effect, the proposed toxin-antidote system would only be active/expressed when

inherited from NIC59 females (specifically, the toxin-antidote system would be active in (N); N/J F1 females, but not in (J); N/J F1 females).

3.3.3 *Is the NIC59 cytoplasmic incompatibility locus encoded in the mitochondrial genome?*

In the (J); N/J female x JU1825 male backcross experiments, we show that individuals homozygous for JU1825 alleles on chromosome III are inviable even in the absence of NIC59 mtDNA. This experiment casts doubt on whether the NIC59 mitochondrial genome encodes the NIC59 cytoplasmic incompatibility locus.

However, another potential explanation consistent with our data is that JU1825 mtDNA is naturally heteroplasmic and contains JU1825-like and NIC59-like alleles. The NIC59-like mtDNA is incompatible with JU1825 chromosome III and is naturally kept at a low frequency within JU1825. However, the NIC59-like mtDNA can increase in frequency in (J); N/J F1 females and eventually cause incompatibilities when backcrossed to JU1825 males. However, the NIC59-like mtDNA does not possess all the SNPs of NIC59 mtDNA, and therefore our PCR assay would fail to detect it. Although it seems unlikely that a NIC59-like mtDNA could persist in a JU1825 nuclear background, there are examples of deleterious heteroplasmic mtDNA deletions in *C. elegans* and *C. briggsae* persisting and even increasing in frequency after many generations of laboratory propagation (Clark *et al.* 2012; Gitschlag *et al.* 2016).

3.4 MATERIALS AND METHODS

See Chapter 4 Materials and Methods for details on the generation of the JU2079 genome, calling fixed SNPs between JU1825 and NIC59, and calculating NIC59 allele frequencies in F2 bulk sequencing and introgression line experiments.

3.4.1 *Bulk collection of viable F2 populations*

We set-up 10 cross plates of 10 JU1825 L4 female x 10 NIC59 L4 male and vice versa to generate (J); N/J F1 females and (N); N/J F1 females, respectively. We then set-up 17 cross plates of 20 (J); N/J F1 female x 20 NIC59 male and 20 (N); N/J F1 female x 20 JU1825 male. The plates were placed at 25° overnight, allowing the L4's to mature to adulthood and begin mating. The next day, adults were permanently removed from the plates. The plates were then placed at 25°overnight. The next day, viable F2 worms were washed off all the plates using 1x M9 into a 15 ml conical tube, washed twice to remove contaminating bacteria and then frozen as an approximately 200 µl pellet of worms at -80°. The worms were then lysed and genomic DNA was extracted using a Gentra Puregene kit. Purified genomic DNA was then submitted to the Fred Hutchinson Genomics core for Illumina library preparation and sequencing on a HiSeq2000 with 100-bp PE reads.

3.4.2 *Generating introgression lines and their sequencing libraries.*

Each introgression line was founded by single F2 females derived from crossing a population of (N); N/J F1 females to JU1825 males. Each F2 female was independently backcrossed to 1-3 JU1825 males. Single females were picked each generation and backcrossed to 1-3 JU1825 males until the F15 generation. At the F15 generation, siblings mated for at least five generations, however some underwent sib-mating for considerable more generations. However, almost all introgression lines were purposely homozygosed for NIC59 alleles at the oPL78+79 locus (JU2079 scaffold2522) on chromosome III, which might not be tightly linked to the nuclear incompatibility locus. Therefore, in these strains, we only require at least 50% NIC59

allele frequencies over the nuclear incompatibility locus. However some strains were not homozygosed for NIC59 alleles at the oPL78+79 locus (XZ1703, XZ1783 and XZ1712_rec53) and therefore we require that the nuclear incompatibility locus have an ~100% NIC59 frequency.

Genomic preps were made from large populations of worms from each introgression line. Illumina sequencing libraries were prepared by shearing 1µg genomic DNA into 300-500 bp fragments in a Bioruptor. Fragmented DNA was cleaned with a genomic DNA concentrator-10 column and eluted in 40 µl of 10 mM Tris-HCl pH 8.0. 18.5 µl of cleaned fragmented genomic DNA underwent end-repair (18.5 µl fragmented genomic DNA, 2.5 µl 10x T4 DNA ligase Buffer w/10 mM ATP (NEB cat #B0202S), 1 µl 10 mM each dNTPs (Invitrogen), 1 µl T4 DNA polymerase (NEB, 3000 U/ml), 1 µl T4 PNK (NEB, 10000 U/ml), 1 µl Roche Standard Taq DNA Polymerase (5000 U/ml, cat#11 146 165 001) for 20 min at 25°, then 20 min at 72°. Illumina adapters were then ligated to the end-repaired DNA (25 µl end-repair reaction, 1 µl 25 µM pre-annealed, forked Illumina TruSeq adaptors, 1 µl Roche Rapid T4 DNA ligase (5000 U/ml, cat# 11 635 379 001)) for 15 mins at 25°. Ligated DNA was then cleaned using a Zymo Clean and Concentrator-5 kit, and eluted in 10 µl 10 mM Tris-HCl pH 8.0. Clean and ligated DNA then underwent three separate minimal PCR amplification reactions (15 cycles) to add the universal forward and index primers (2 µl clean and ligated DNA, 10 µl Roche Hi-fi Buffer w/o $MgCl_2$, 1 µl 10 mM dNTP each, 5 µl $MgCl_2$, 1 µl 50 µM multiplexing PCR primer 1, 1 µl 50 µM Truseq indexing primer, 1 µl Roche High Fidelity Plus enzyme, 29 µl MilliQ dH_2O). All three separate PCR products were then pooled together and cleaned up using a Zymo DNA Clean and Concentrator-25 kit and eluted in 30 µl of Tris-HCl pH 8.0. PCR amplified samples then underwent an ampure bead clean-up to get rid of adapter-adapter PCR products. Each sequencing library was added in equal molar amounts to a 3 nM pool, which was submitted to the Fred

Hutchinson Genomics core for sequencing on a Illumina HiSeq2000 with 50-bp paired-end reads.

3.4.3 *Scoring viability*

To score the viability of the offspring of a cross, 10-15 L4 females and males were placed on a plate and placed at 25°overnight. Overnight, the worms mature into adults and begin mating and laying embryos. The next day, the now adult worms are move onto a new plate and allowed to lay eggs for 2 hours at 25°. After 2 hours, the parents are permanently removed from the plates and the number of embryos they laid are immediately counted. The next day, we quantified the number of embryos that failed to hatch. We define "percent embryonic lethality" as the number of unhatched embryos divided by the total number of embryos laid. The next day, we placed the plates at 4° for 1 hour and then quantified the number of healthy L4 larvae and adults were on the plate. We define "percent viable progeny" as the total number of healthy L4 larvae and adults divided by the total number of embryos laid. We find that the number of unhatched embryos and the number of viable L4/adults do not always add up to the total number of embryos laid. This implies that some fraction of offspring survived embryogenesis but have not developed to the L4 or adult stage at the time of scoring, either because of slow larval growth or inviability. We call this difference the number of "missing larvae" and define "percent missing larvae" as the number of missing larvae divided by the total number of embryos laid.

3.4.4 *Gene annotation of JU2079 scaffold19*

The JU2079 genome was annotated by Lewis Stevens and Mark Blaxter using MAKER2 (Holt and Yandell 2011) and JU2079 RNAseq data. JU2079 scaffold19 annotated with coding

sequence was downloaded from caenorhabditis.org. The coding sequences were manually translated into amino acids.

3.4.5 *Phylogenetic analysis*

We constructed two maximum-likelihood phylogenies using the amino-acid sequences of all F-box and nuclear hormone receptor genes in the candidate region. JU2079 coding sequence annotations were obtained from caenorhabditis.org, and NIC59 coding sequences were inferred based on alignments to homologous JU2079 coding sequence(s). The amino-acid sequences were aligned with MUSCLE (Edgar 2004) using default settings. A maximum likelihood phylogeny was made using the PhyML function in Ugene (Okonechnikov *et al.* 2012). We used an LG substitution model and tree searching was performed Nearest Neighbor Interchanges (NNIs). Branch support was determined using the aLRT likelihood method.

3.4.6 *Making the NIC59 fosmid library*

9 µg of NIC59 genomic DNA was run on a 1% LMP ultrapure gel 17 hours overnight on a pulse field gel box (cooling at 14°, pump=65, 170V, initial A=1 final A=6). Genomic DNA fragments around 30-40 kb were cut out and gel purified by melting the gel slice in an Eppendorf tube at 70° for 15 min and spinning down at max speed. 2 µl of agarase (Promega) was mixed into the supernatant and incubated at 47° for 1 hour. The sample was then put on ice for 5 min and spun down at max speed for 20 min. The supernatant containing NIC59 genomic DNA was then cleaned and concentrated using Ampure beads, and eluted in 60 µl of 10 mM Tris-HCl pH 8.0 (11 ng/µl). The entire sample underwent end-it end-repair (60 µl sample, 5 µl MilliQ dH_2O, 10 µl 2.5 mM dNTP mix, 10 µl 10 mM ATP, 10 µl 10X End-It Buffer, 5 µl End-IT Enzyme Mix) at 25° for 45 min, then 70° for 15 min. The end-repaired sample then was cleaned and

concentrated using Ampure beads, and eluted in 8 µl of 10 mM Tris-HCl pH 8.0. The cleaned end-repair sample was then ligated to the pCC1FOS linear vector (8 µl end-repaired sample, 1.2 µl NEB 10X T4 DNA ligase Buffer w/ATP, 1 µl pCC1FOS, 1 µl NEB T4 DNA Ligase and 0.8 µl MilliQ dH_2O) at 16° overnight, then at 70° for 10 min. 10 µl of ligated sample was then mixed with 25 µl of phage packaging extract and incubated at 30° for 90 min. Then another 25 µl of packaging extract was added and the sample was incubated a second time at 30° for 90 min. Then the phage-packaged sample was diluted in 440 µl of phage dilution buffer (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$). We determined the titer of the diluted phage-packaged sample by inoculating 100 µl of EPI-300 bacteria in 10 mM $MgSO_4$ (OD 0.75A) with either 0.5 µl, 1.0 µl, 3.0 µl or 5.0 µl of diluted phage-packaged sample and quantifying the number of colonies each produced when plated on an LB-chloramphenicol plate. We extrapolated that there are ~15,000 colony forming units (clones) in our diluted phage-packaged sample. Given an average insert size of 30 kb, then we would on average have 6x coverage of the 75-Mb *C. nouraguensis* genome.

We planned to distribute 15,000 clones across three 96-deep-well plates, with each well having ~50 clones. All of the 500 µl diluted phage-packaged sample (should give 15,000 clones) was added to 30 ml of EPI-300 bacteria in 10 mM $MgSO_4$ (OD 0.75A), and placed in a 37° incubator without shaking for 1 hour. We then added 30 ml LB to the sample and distributed the now 60 ml across all three 96 deep-well plates (200 µl/well). We then added 800 µl LB to each well, covered the plates and incubated them at 37° for 45 mins with shaking. Afterwards, we added 500 µl LB + 0.75 µl 25 mg/mL chloramphenicol to each well and incubated all three 96 deep-well plates at 37° overnight with shaking at 180 rpm. The next day, 400 µl of each wells bacterial overnights was moved into its corresponding well of a new 96-deep well plate. 625 µl

of 40% glycerol was added to each of these new wells and mixed by pipetting. The three 96 deep-well plates were then frozen immediately at -80°.

3.4.7 *Colony hybridization*

The titer of an overnight growth of a NIC59 fosmid library pool or subpool (in LB + chloramphenicol) was determined by plating serial dilutions (20 uL of overnight in 180 uL of LB + chloramphenicol). Based on the titers, we made two plates with approximately 500 colonies each. The plates were cooled at 4° for 30 min. A nylon membrane was then placed on the surface of each plate and left for 2 min. We then removed the nylon membrane and placed it colony-side up in 1 ml of denaturation solution (0.5 M NaOH, 1.5 M NaCl) for 15 mins, in 1 ml of neutralization solution (1.5 M NaCl, 1.0 M Tris-HCl pH 7.4) for 15 mins and then in 1 mL of 2x SSC (0.3 M NaCl, 30 mM sodium citrate, pH 7.0) for 10 mins. The DNA was then cross-linked to the nylon membrane using shortwave UV radiation for 1 min. Cell debris were then removed from the membrane by placing 500 µl 2X SCC + 50 µl Proteinase K (>600 U/ml) directly onto the membrane and incubating at 37° for 1 hour. The nylon membrane was placed in a 50 mL falcon tube with 10 ml of preheated DIG Easy Hyb buffer for 30 mins at the hybridization temperature (44° for probe oPL431+432) with rotating (prehybridization). Meanwhile, 10 µl of dsDNA DIG probe (DIG-PCR product oPL431+432) was added to 40 µl of MilliQ dH_2O and denatured at 97° for 5 min and then immediately placed on ice. All 50 µl of denatured DIG probed was then added to 6 mls of preheated DIG Easy Hyb buffer. The prehybridization DIG Easy Hyb buffer was then discarded from the falcon tube containing the nylon membrane and 6mls of DIG Easy Hyb buffer + probe was added. Probes hybridized overnight at 42° with rotating. The following day, the DIG Easy Hyb + probe were discarded and the nylon membrane was washed twice with 6 ml of low stringency buffer (2x SCC + 0.1% SDS) for 8 mins at room

temperature, once with 6 ml of high stringency buffer (0.5x SCC + 0.1% SDS) for 15 mins at 65°, and then once with 10 ml of washing buffer for 8 mins at room temperature. The washing buffer was discarded, and 20 ml of blocking solution was added to the membrane and incubated for 1 hour at room temperature while rotating. Afterwards, the blocking solution was discarded and 20 mls of antibody solution (2 µl anti-DIG antibody + 20 ml blocking solution) was added to the membrane and incubated for 45 min at room temperature while rotating. The antibody solution was discarded, and the membrane was washed twice with 20 ml of washing buffer for 15 min each. The washing buffer was discarded, and 20 ml of detection buffer was added to the membrane and incubated for 3 min at room temperature. The nylon membrane was then removed from the falcon tube and placed colony-side down into 500 µl pool of a 1:100 dilution of 25 mM CSPD (in detection buffer) on top of a transparency film. The nylon membrane sealed with another piece of transparency film, placed at 37° for 10 min, and then imaged.

3.5 ACKNOWLEDGEMENTS

We thank Marie-Anne Félix and Christian Braendle for providing all the strains of *C. nouraguensis* used in this study; Janet Young for assembling the JU2079 genome and calculating NIC59 allele frequencies in bulk F2 and introgression line samples; Luke Noble and Matthew V. Rockman for assembling and sharing the *C. becei* genome; Mark Blaxter for sequencing JU2079 (DNA and RNA); Antoine Molaro, Michelle Hayes and Harmit Malik for advice and resources for generating Illumina TruSeq libraries. Ruolan Qiu and Jay Shendure for advice and resources for generating the NIC59 fosmid library; Maulik Patel for purifying genomic DNA from viable F2 populations; and the Fred Hutchinson Genomics core for Illumina library preparation and sequencing.

3.6 FIGURES

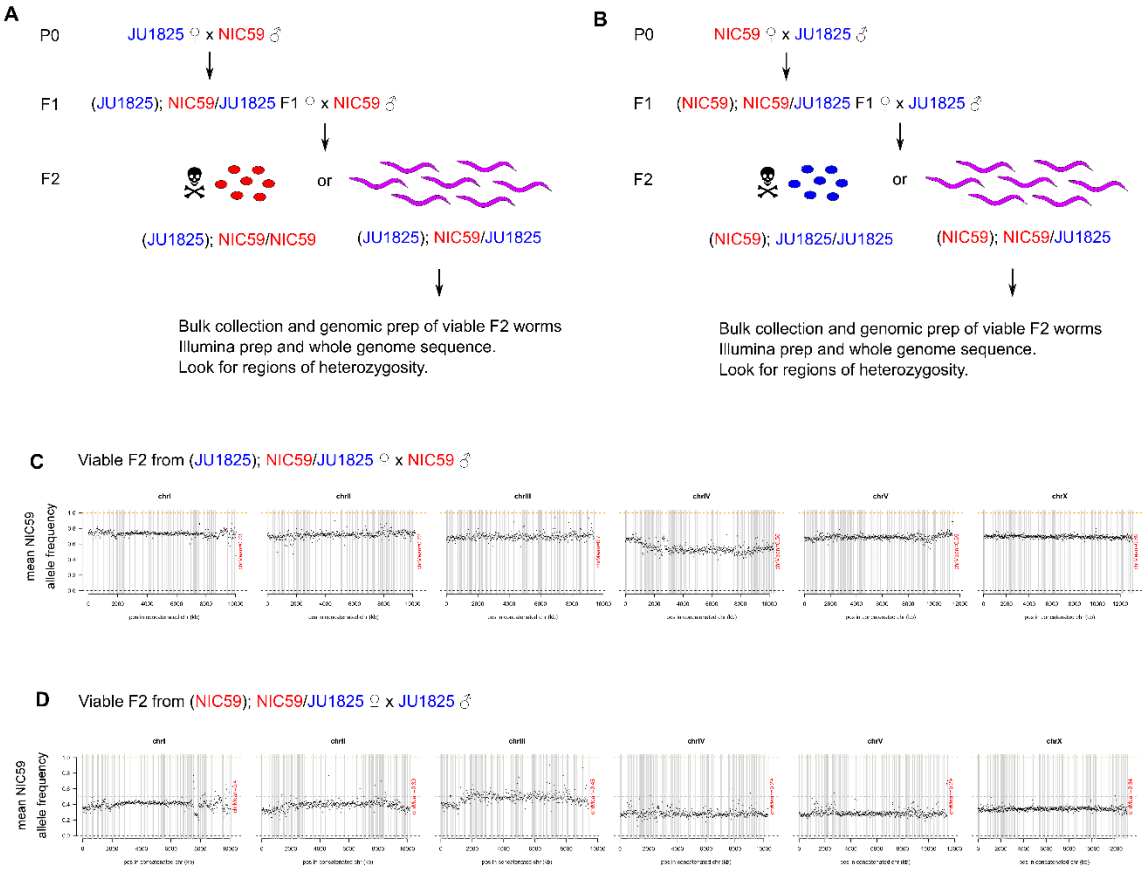


Figure 3.1. Bulk sequencing of viable F2 offspring reveal genetically distinct nuclear incompatibility loci.

(A) A diagram illustrating our hypothesis that the JU1825 cytoplasm is incompatible with a single nuclear locus homozygous for NIC59 alleles and causes F2 death (dead embryos symbolized by red ovals). Presumably F2 individuals that are heterozygous at this locus are viable (purple worms), therefore collecting and sequencing a population of these can help map the NIC59 nuclear incompatibility locus. **(B)** A diagram illustrating our hypothesis that the NIC59 cytoplasm is incompatible with a single nuclear locus homozygous for JU1825 alleles and causes F2 death (dead embryos symbolized by blue ovals). Presumably F2 individuals that

are heterozygous at this locus are viable (purple worms), therefore collecting and sequencing a population of these can help map the JU1825 nuclear incompatibility locus. **(C)** Whole-genome sequencing data from a population of viable F2 from the (JU1825); NIC59/JU1825 female x NIC59 male backcross. Each plot represents one of the six chromosomes of the *C. nouraguensis* assembly (see Chapter 4 Materials and Methods). Each point represents the average NIC59 SNP frequency in 20-kb windows ordered along the chromosome. The *C. nouraguensis* assembly is fragmented into smaller scaffolds (shown as grey lines) that are ordered based on synteny to *C. becei* (see Chapter 4 Materials and Methods). **(D)** Whole-genome sequencing data from a population of viable F2 from the (NIC59); NIC59/JU1825 female x JU1825 male backcross.

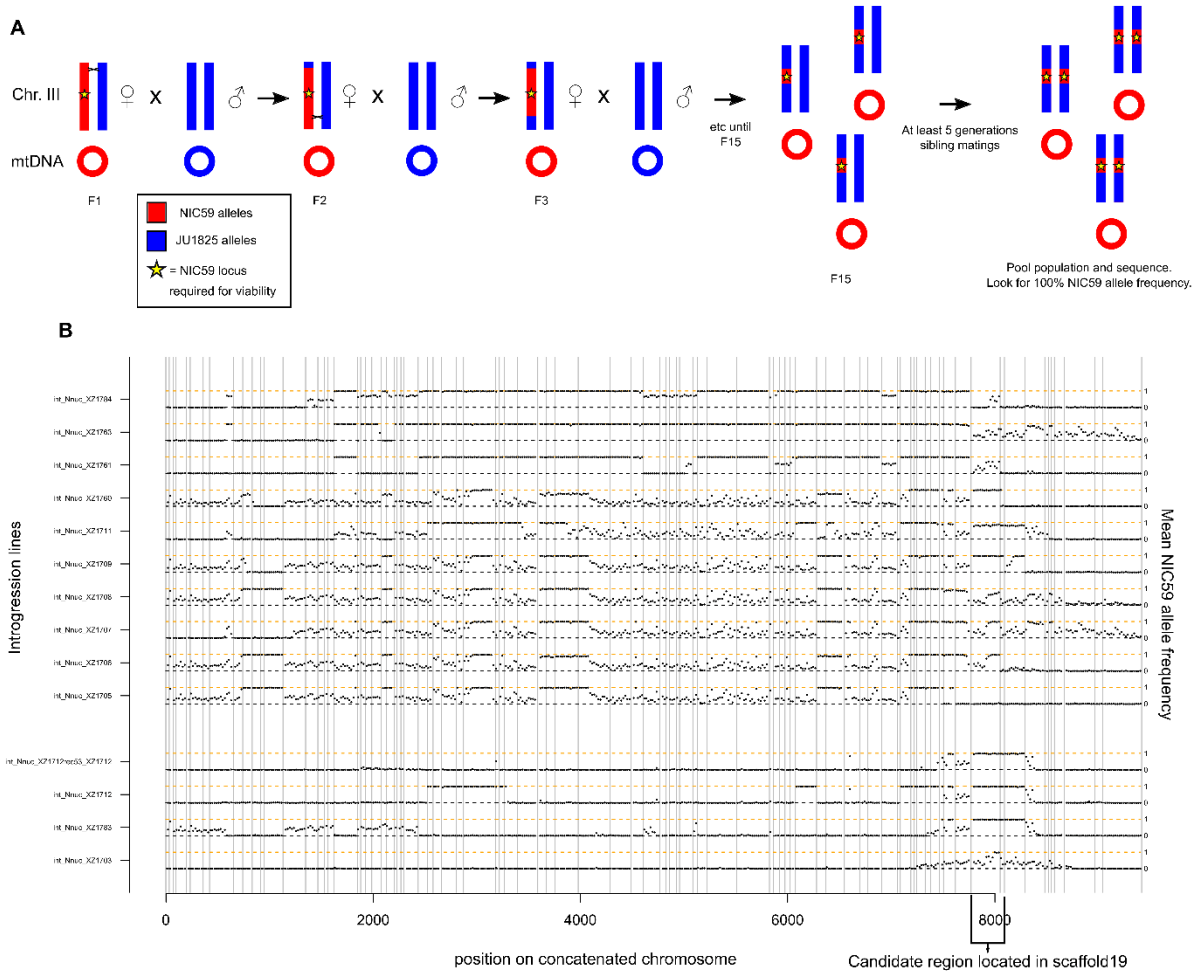


Figure 3.2. Sequencing introgression lines maps the JU1825 incompatibility locus to JU2079 scaffold19.

(A) A schematic showing how introgression lines can help finely map the JU1825 nuclear incompatibility locus. Heterozygous NIC59/JU1825 F1 females with a NIC59 mitochondrial genotype are backcrossed to JU1825 males. Surviving progeny will be heterozygous at the NIC59 locus required for viability (star on chromosome III). For each introgression line, only one surviving female is then mated to JU1825 males each backcross generation for 15 generations. Recombination occurring during each backcross generation will increasingly make loci unlinked to the incompatibility locus on chromosome III homozygous for JU1825 alleles. At the F15 generation, siblings were mated for at least five generations to homozygose the NIC59

locus required for viability. A large population of the introgression lines were then sequenced and scanned for regions on chromosome III that are homozygous for NIC59 alleles. **(B)** Plots showing NIC59 allele frequencies in 20-kb windows across chromosome III for various introgression lines. The x-axis indicates physical length (in kb) of JU2079 chromosome III and the y-axis indicates NIC59 allele frequency (0.0-1.0) The gray vertical lines indicate breaks between JU2079 scaffolds assigned to chromosome III. Introgression line XZ1703 is homozygous for NIC59 alleles only on JU2079 scaffold19.

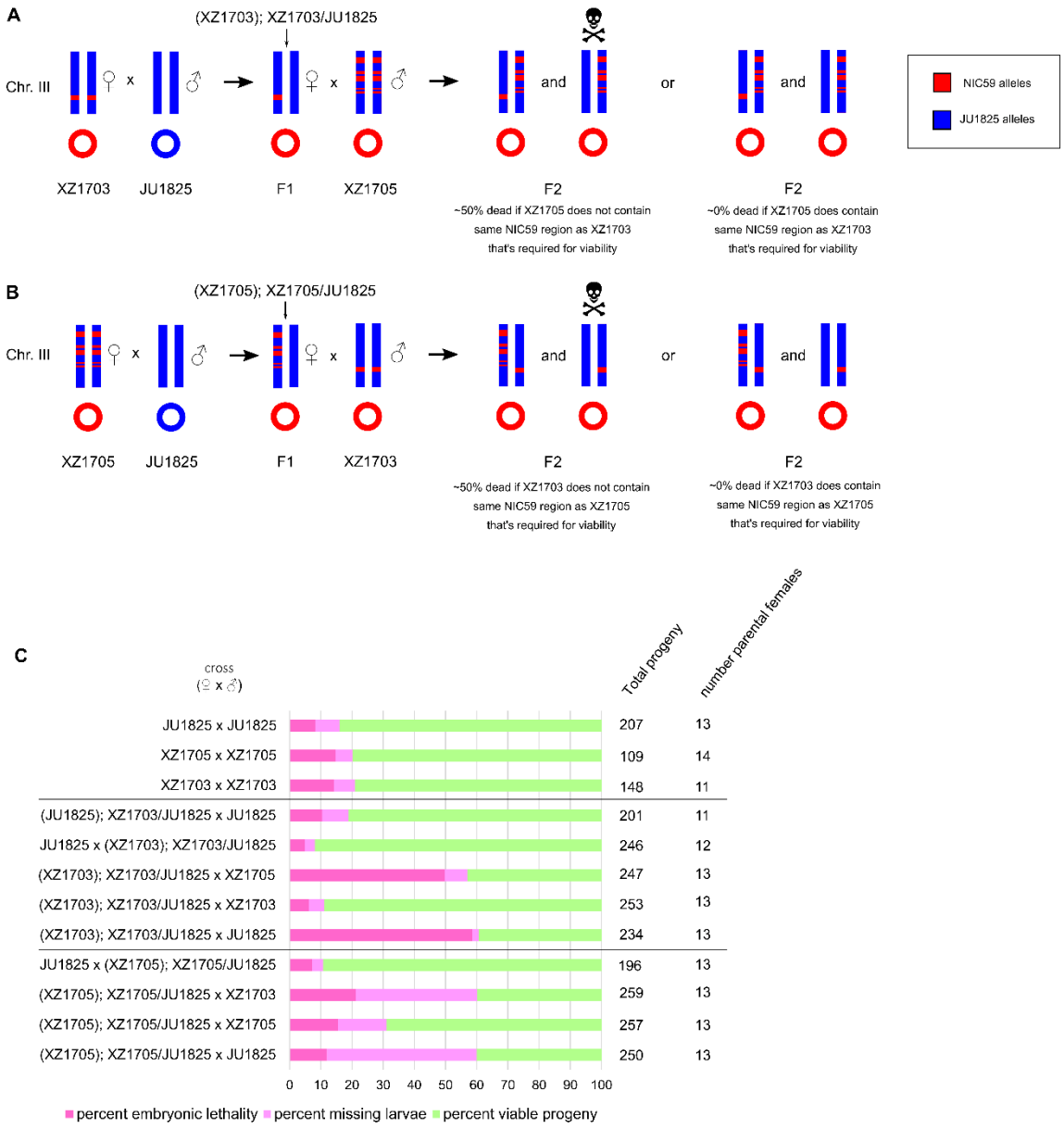
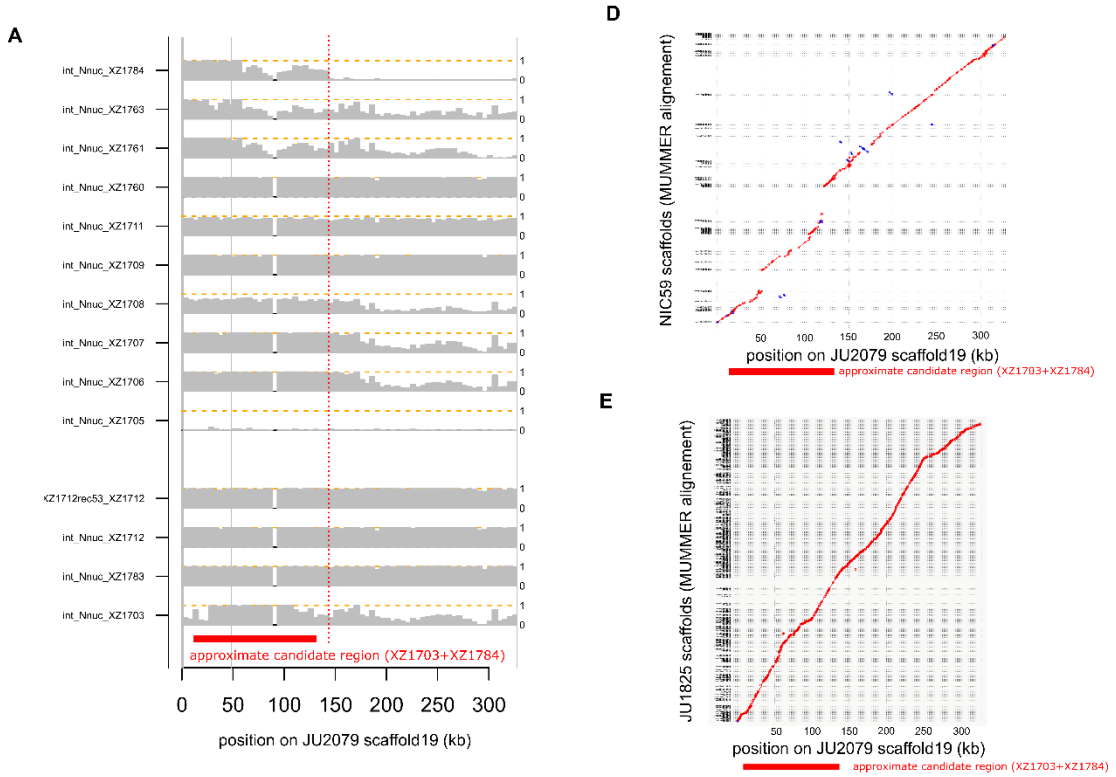
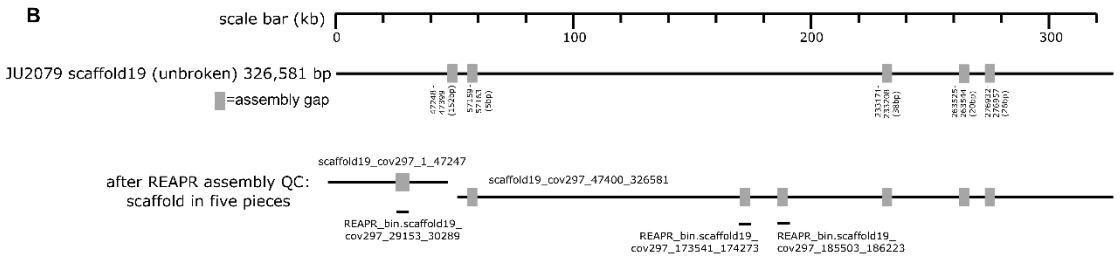


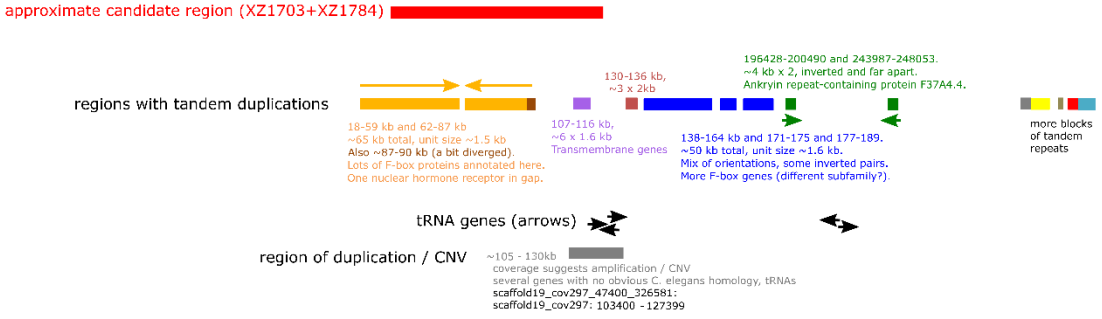
Figure 3.3. Introgression line XZ1705 cannot rescue F2 inviability caused by the XZ1703 cytoplasmic – JU1825 nuclear incompatibility.

(A) A schematic illustrating how we can test whether XZ1705 has NIC59 loci on Chr. III (red dashes that roughly map to NIC59 regions in Figure 3.2) that can rescue the XZ1703 cytoplasmic – JU1825 nuclear incompatibility. **(B)** A schematic illustrating how we can test whether XZ1703 has NIC59 nuclear loci that can rescue the XZ1705 cytoplasmic – JU1825 nuclear incompatibility. **(C)** A bar graph quantifying the viability (x-axis) of various crosses (y-axis).

The number of parental females used in each cross and the number of offspring they had are indicated. Crossing F1 males to JU1825 females are meant as controls for high F2 viability.







C

Number of F2 with either a NIC59/JU2079 or JU2079/JU2079 genotype at Chr. III (oPL78+79)

cross (female x male)	NIC59/JU2079 Heterozygote	JU2079 Homozygote
(NIC59); NIC59/JU2079 F1 x JU2079	38	1
JU2079 x (NIC59); NIC59/JU2079 F1	20	20

Figure 3.4. The JU1825 nuclear incompatibility locus maps to an ~100 kb region on JU2079 scaffold19.

(A) Plots showing NIC59 allele frequencies in 5-kb windows across JU2079 scaffold19 for various introgression lines. The x-axis indicates the physical length of the scaffolds in kb, while the y-axis indicates the NIC59 allele frequency (0.0-1.0). The vertical red dotted line demarcates the right boundary of the candidate region because introgression line XZ1784 is homozygous for JU1825 alleles to the right of it. To the left of the dashed line is a thick horizontal red line that underlines positions in XZ1703 that have 1.0 NIC59 allele frequencies, and therefore is our JU1825 nuclear incompatibility locus candidate region. **(B)** A diagram depicting some features of JU2079 scaffold19. JU2079 scaffold19 was initially ~326 kb in length, but then broken into five pieces by REAPR (Hunt *et al.* 2013), which is used to break apart misassembled scaffolds. The candidate region is indicated by the thick red horizontal line. Below that are horizontal lines indicating regions containing tandem duplications. Each color corresponds to a group of related tandem duplications. Beneath them are brief descriptions of the regions they span, their total size, the length of their repeating unit and what kinds of genes are annotated within them (if applicable). Single tRNA genes are indicated by black arrows. The thick gray horizontal line marks a region where there is large copy-number variation between NIC59 and JU1825. **(C)** A table showing that the NIC59 cytoplasm is incompatible with JU2079 chromosome III. 38/39 viable F2 derived from crossing (NIC59); NIC59/JU2079 F1 females to JU2079 males are heterozygous NIC59/JU2079 at a chromosome III marker, indicating that F2 that are homozygous are inviable. As a control, we show that crossing JU2079 females to (NIC59); NIC59/JU2079 F1 males shows no selection against homozygosity (equal proportions of NIC59/JU2079 heterozygotes and JU2079/JU2079 homozygotes). **(D)** A MUMMER plot

showing NIC59 scaffolds (y-axis) aligning to JU2079 scaffold19 (x-axis). The black horizontal and dotted lines indicate breaks between NIC59 scaffolds. Red lines indicate forward alignments between the two assemblies, whereas blue lines indicate inverted alignments. The approximate position of the candidate region is indicated by the thick red horizontal line. **(E)** A MUMMER plot showing JU1825 scaffolds (y-axis) aligning to JU2079 scaffold19 (x-axis). The black horizontal lines indicate breaks between NIC59 scaffolds, respectively. Red lines indicate forward alignments between the two assemblies, whereas blue lines indicate inverted alignments. For a list of all NIC59 scaffolds that map to the candidate region, see Table 3.2. The approximate position of the candidate region is indicated by the thick red horizontal line.

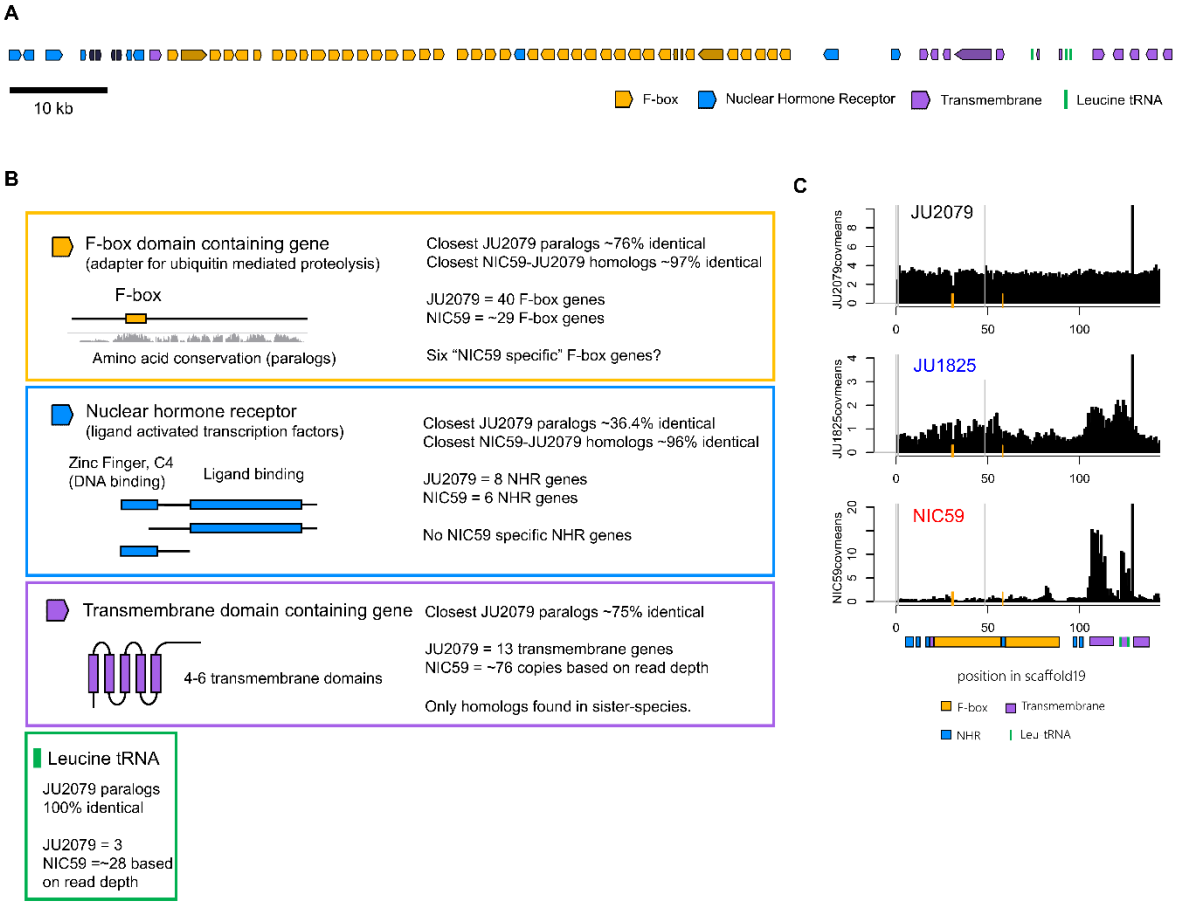


Figure 3.5. The nuclear candidate region contains F-box, transmembrane, nuclear hormone receptor and tRNA genes.

(A) A cartoon representation of the genes annotated in the candidate region on JU2079 scaffold19, with each arrow representing a single gene. Darker versions of a color are predicted to be misannotations that either fuse multiple genes or fragment a single gene. **(B)** An infographic describing the general features of the four types of genes found in the candidate region. Each infographic describes the amino acid identity between the closest paralogs in the JU2079 assembly as well as between the closest NIC59-JU2079 homologs. Each infographic also compares the number of genes NIC59 and JU2079 encode in the candidate region. **(C)**

Graphs showing JU2079, JU1825 and NIC59 read coverage across the candidate region in 2-kb windows. The y-axis measures relative read coverage while the x-axis indicates the physical length of JU2079 scaffold19 in the candidate region. Below the graphs are colored boxes indicating roughly where F-box, transmembrane, nuclear hormone receptor and leucine-tRNA genes lie.

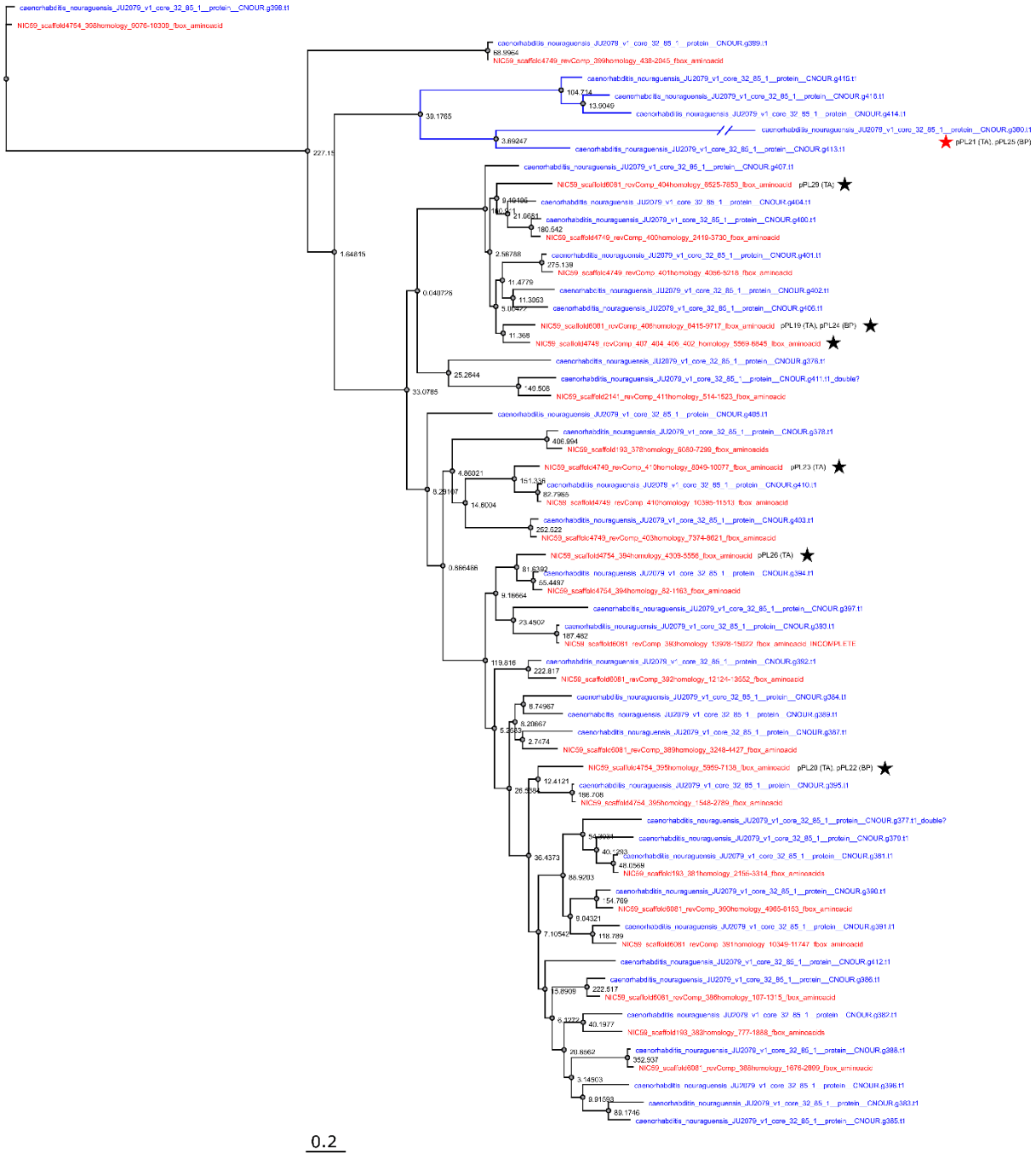


Figure 3.6. NIC59 has six F-box genes with no clear JU2079 homologs.

A maximum-likelihood phylogeny of all NIC59 (red) and JU2079 (blue) F-box proteins in the candidate region. The six NIC59 F-box genes without immediate JU2079 homologs are marked

by black stars. The plasmids in which these NIC59-specific F-box genes have been cloned into are also indicated (TA clone and/or BP clone). The red star symbolizes a NIC59 F-box gene whose immediate homolog is likely JU2079 g380.t1 (gene 380, transcript 1, gene names generated by caenorhabditis.org), but has an early stop codon. Each NIC59 F-box name is a description of which NIC59 scaffold it is found in and what JU2079 F-box gene it is most closely related to. The scale bar indicates 0.2 average substitutions per site.

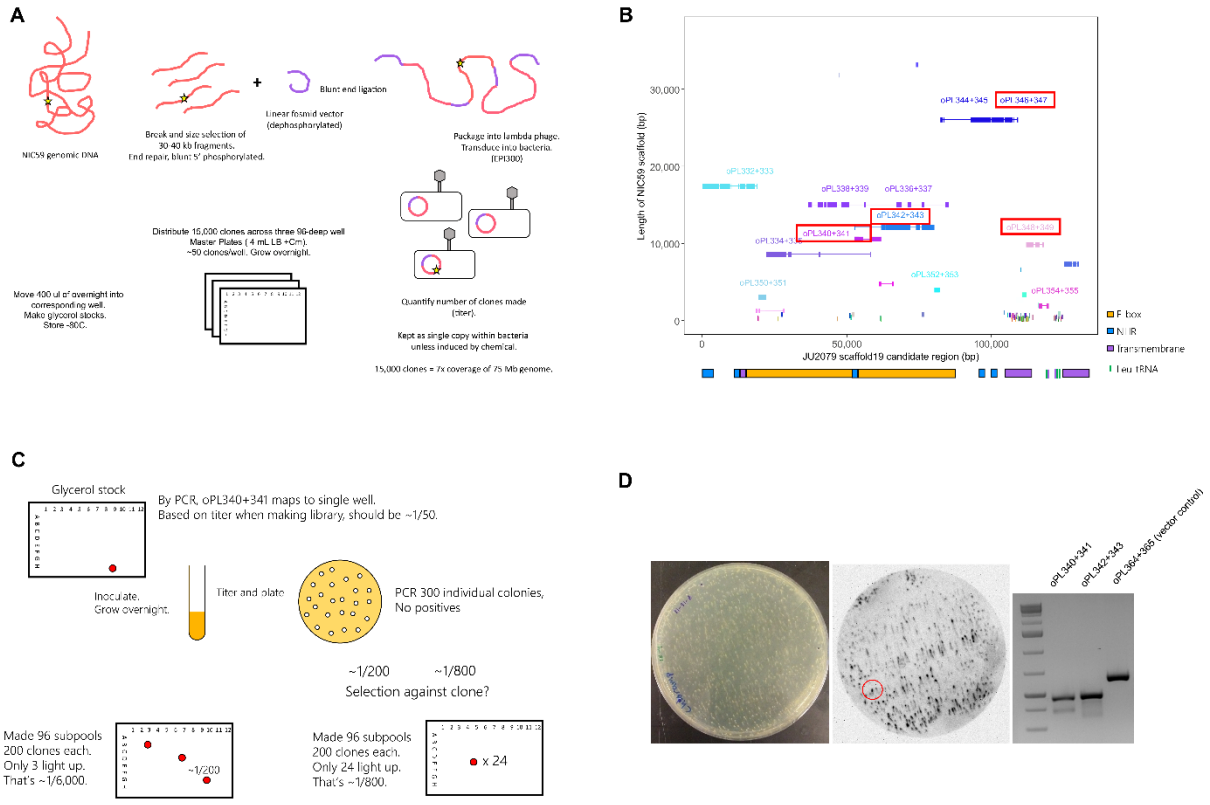


Figure 3.7. Generating and screening the NIC59 fosmid library.

(A) A diagram showing how the NIC59 fosmid library was generated (see Materials and Methods). **(B)** A graph showing all the NIC59 scaffolds (different colored horizontal lines) that align to the JU2079 candidate region using MUMMER. The y-axis indicates the length of NIC59 scaffolds, while the x-axis indicates where they map along the physical length of the JU2079 candidate region. The thick lines represent regions of alignment between NIC59 and JU2079, while thin lines just connect different aligning regions of the same NIC59 scaffold. The sets of primers used to screen the NIC59 fosmid library are shown above the NIC59 scaffolds they map to. The primers that gave positive signals when screening the fosmid library are boxed in red. Below the graph are colored boxes indicating roughly where F-box, transmembrane, nuclear

hormone receptor and leucine-tRNA genes lie. **(C)** A set of experiments that demonstrate a NIC59 fosmid of interest (oPL340+341 positive, see Figure 3.7B) is selected against when grown in a pool containing other clones. **(D)** The first image is of an LB plate used to lift colonies for DIG-probe colony hybridization (bacteria from a subpool of Plate1Well9H of the NIC59 fosmid library). **(E)** The second image shows a blot of the lifted colonies using an anti-DIG antibody. The third image is of a DNA gel showing that one of the colonies (indicated by the red circle in the second image) is PCR positive for two sets of screening primers.

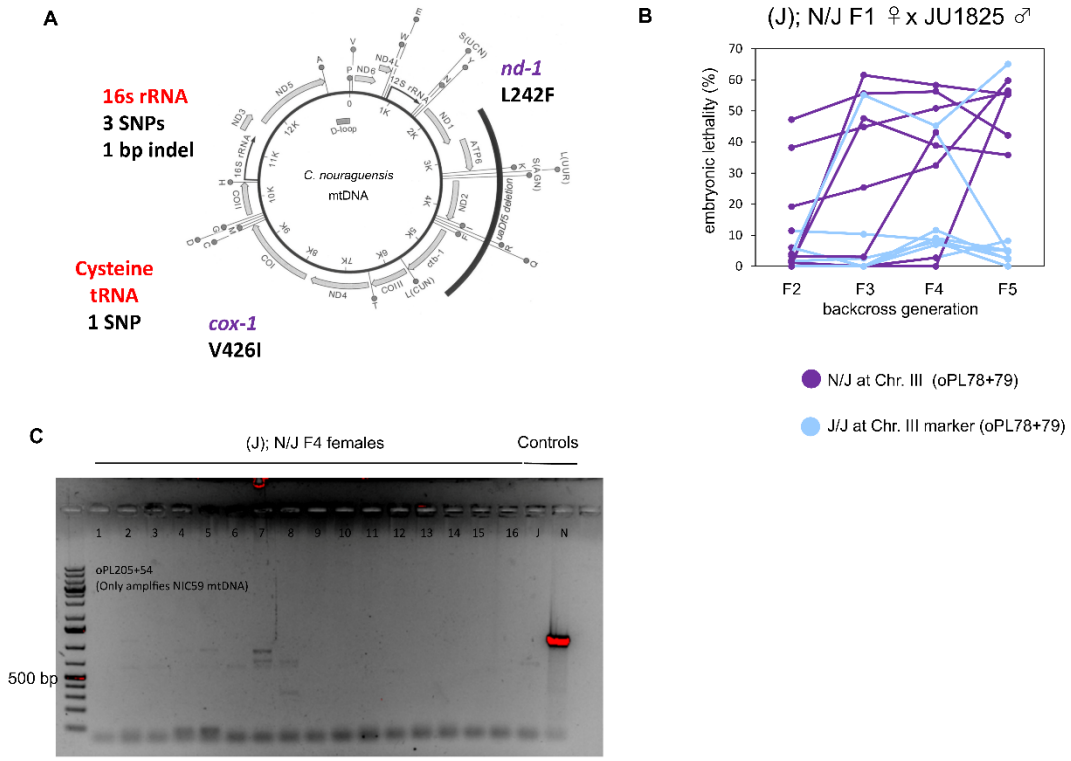


Figure 3.8. Mapping the NIC59 cytoplasmic incompatibility locus.

(A) A diagram of the *C. nouraguensis* mitochondrial genome (adapted from (Lemire 2005)). Grey boxed arrows represent protein coding genes, thin black arrows represent rRNA genes and dark circles represent tRNA genes. Potential functional differences between the NIC59 and JU1825 mitochondrial genomes are noted, including non-synonymous SNPs and SNPs in tRNA and rRNA genes. **(B)** A graph quantifying the percent embryonic lethality of different (J); N/J female x JU1825 male backcross lineages over multiple generations. Each line represents a single backcross lineage, and its color indicates whether the female that gave rise to that generation was heterozygous NIC59/JU1825 (purple) or homozygous JU1825 (blue) for a chromosome III locus. **(C)** A DNA gel showing that none of the F4 females in the backcross lineages in Figure 3.8B were positive for a NIC59 mtDNA-specific PCR reaction.

3.7 TABLES

Table 3.1. **Primers.**

A list of primers used in this study and their sequences.

Primer name	Sequence (5'-3')
oPL78	TCCTCCAGCATATCCGCTCG
oPL79	CAATTGCACGGAGAGACTGTTCATC
oPL205	CAGCTACTATAGTTATTGCTGTTCCCACC
oPL254	TTTGCGACCCTTCAGGAGTGC
oPL332	CAGTTCTAGCTTCGAATTCTGG
oPL333	TCTCTTATCAGCTCTTATCACCTG
oPL334	GAATGCGATTTTCCGTCGAAC
oPL335	AGTATGGCTACACGCAACTCTG
oPL336	ATAGTCAAGTATGCCGAATCACCC
oPL337	TTCGAATGAACAACAGCACG
oPL338	GCAAGTAATCAAGTATATGCCC
oPL339	AGACTTTATCGAATGATAACG
oPL340	ATCAGAGTATCGAGATGGTCCG
oPL341	ATCGGAGGAGATGAAGTTTCC
oPL342	GATTCATTCTCACTGCTGATCC
oPL343	TACAGGATGAGATCCGTGTC
oPL344	AATCCGTGTCATTCAATTTCG
oPL345	CTCCTTATCACTTATCACTCACCC
oPL346	AGCATGTTCGATGTGATTTGG
oPL347	ATTGATTTTCCGACTTCCGC
oPL348	TGATTGTATTCTTGTCGGCCAG
oPL349	TTTGCAGACTTCTCGATCGC
oPL354	GCAAGGCATCTTGTCTCTCG
oPL355	GTCTAGCGCTAATTCCAGTGC
oPL431	GAGTGAGCGAGGAAGCACCAGG
oPL432	CATCGCCGGCATCCTCTTCAGG

Table 3.2. **NIC59 scaffolds mapping to candidate region using MUMMER.**

Each region on a NIC59 scaffold that maps the candidate region is noted.

NIC59 scaffold that aligns to JU2079 scaffold19 candidate region	start alignment in JU2079 scaffold19 (bp)	end alignment in JU2079 scaffold19 (bp)	start alignment in NIC59 scaffold (bp)	end alignment in NIC59 scaffold (bp)	percent identity	length of NIC59 scaffold (bp)	% NIC59 scaffold aligned
scaffold5957	352	5,479	17,261	12,166	90.38	17,327	29.41
scaffold5957	5,296	5,904	12,164	11,555	92.18	17,327	3.52
scaffold5957	6,205	9,849	11,839	8,312	86.12	17,327	20.36
scaffold5957	10,314	10,540	6,594	6,820	91.74	17,327	1.31
scaffold5957	12,595	12,864	6,860	6,590	91.94	17,327	1.56
scaffold5957	13,219	15,090	6,256	4,357	89.37	17,327	10.97
scaffold5957	15,552	16,541	4,283	3,311	89.03	17,327	5.62
scaffold5957	16,323	18,256	1,608	3,522	92.01	17,327	11.05
scaffold5957	18,044	18,331	1,826	1,529	89.97	17,327	1.72
scaffold5957	18,360	18,518	885	728	86.88	17,327	0.91
scaffold5990	18,698	19,232	719	1,253	94.02	1,253	42.7
scaffold5957	18,770	19,232	460	1	92.89	17,327	2.65
C233380	19,166	19,493	1	328	100	328	100
C209282	19,427	19,657	231	1	95.67	231	100
scaffold4543	19,427	19,662	1	236	94.92	2,975	7.93
C169838	19,591	19,741	1	151	99.34	151	100
scaffold4543	19,664	22,159	440	2,904	87.72	2,975	82.86
scaffold5990	19,675	20,029	1	359	88.58	1,253	28.65
scaffold193	22,168	24,827	8,527	5,821	89	8,590	31.51
scaffold193	24,656	26,018	3,549	2,179	86.46	8,590	15.96
scaffold193	26,115	29,268	5,514	2,340	94.07	8,590	36.96
C184342	26,241	26,419	1	179	94.97	179	100
C263156	27,285	27,819	1	535	91.04	793	67.47
scaffold5990	28,157	28,491	311	645	89.85	1,253	26.74
scaffold193	29,917	30,900	2,347	1,325	87.5	8,590	11.91
scaffold6081	36,692	37,978	14,946	13,652	90.21	15,022	8.62
scaffold6081	39,869	41,959	13,395	11,316	94.07	15,022	13.85
scaffold193	40,455	41,050	1,364	768	83.55	8,590	6.95
scaffold6081	42,329	43,168	9,220	8,415	82.35	15,022	5.37
scaffold6081	43,354	44,746	9,999	8,599	93.52	15,022	9.33

scaffold6081	44,726	46,763	4,960	2,960	87.74	15,022	13.32
scaffold6081	45,638	46,396	11,364	10,622	80.5	15,022	4.95
C194692	46,764	46,962	199	1	96.48	199	100
scaffold625	46,764	47,247	87,643	87,160	98.76	97,067	0.5
scaffold2903	47,400	47,513	30,900	30,787	95.61	31,817	0.36
scaffold625	47,400	47,512	87,028	86,916	100	97,067	0.12
scaffold6081	48,159	49,223	2,947	1,886	88.65	15,022	7.07
scaffold6081	49,223	50,945	1,723	1	97.97	15,022	11.47
C256744	50,890	51,423	1	553	93.13	608	90.95
C240314	51,429	51,807	1	379	96.57	379	100
C261910	51,809	52,557	749	1	98.26	749	100
scaffold4754	52,557	54,287	4,258	5,971	85.86	10,518	16.3
scaffold4749	52,558	52,970	3,410	3,838	85.08	12,080	3.55
scaffold4754	52,718	55,693	3	2,966	94.84	10,518	28.18
scaffold6081	55,920	56,657	2,807	2,069	80	15,022	4.92
scaffold193	57,667	57,794	1,830	1,703	86.72	8,590	1.49
scaffold193	58,023	58,405	1,464	1,100	80.78	8,590	4.25
scaffold4754	58,541	61,964	6,842	10,310	89.57	10,518	32.98
scaffold36	61,021	61,982	3,327	4,332	88.81	4,753	21.17
scaffold36	61,085	61,983	195	1,086	89.89	4,753	18.77
C233970	61,496	61,823	332	1	91.27	332	100
C161952	61,771	61,905	135	1	98.52	135	100
scaffold4749	61,871	62,379	12,080	11,572	90.73	12,080	4.21
scaffold4749	62,381	63,223	11,453	10,666	90.63	12,080	6.52
scaffold4749	63,412	70,823	10,478	3,085	91.65	12,080	61.21
scaffold36	65,780	66,394	1,816	2,443	81.72	4,753	13.21
scaffold6081	67,253	69,025	5,177	6,940	84.14	15,022	11.74
scaffold4749	70,526	72,095	9,931	8,348	86.91	12,080	13.11
scaffold6081	70,949	72,316	7,349	8,730	83.72	15,022	9.2
scaffold730	73,961	74,845	22,455	23,333	81.86	33,158	2.65
scaffold4749	74,021	75,500	6,620	5,176	84.16	12,080	11.96
scaffold6081	75,838	77,004	5,128	6,296	84.19	15,022	7.78
scaffold4749	75,891	77,271	6,640	5,257	85.05	12,080	11.46
C262392	75,924	76,710	767	1	81.7	767	100
scaffold4749	77,209	80,275	3,642	537	92.31	12,080	25.71
scaffold6099	80,316	82,013	96	1,809	89.3	3,933	43.58
scaffold6099	80,325	82,179	1,763	3,619	92.37	3,933	47.22
scaffold2141	82,357	83,195	25,502	24,656	83.33	26,015	3.26
scaffold2141	83,520	83,743	24,042	23,815	85.59	26,015	0.88

scaffold6081	84,015	85,132	8,885	9,998	79.04	15,022	7.42
scaffold2141	92,734	99,791	19,734	12,707	91.37	26,015	27.02
scaffold2141	99,950	104,283	12,648	8,322	90.52	26,015	16.63
C267346	104,195	104,568	842	460	80.71	980	39.08
scaffold2141	104,528	105,017	5,844	5,354	93.1	26,015	1.89
C267346	104,750	104,881	264	133	87.12	980	13.47
scaffold2141	105,052	106,012	5,461	4,520	84.51	26,015	3.62
C258622	105,384	105,939	550	1	87.43	655	83.97
scaffold2141	106,075	107,238	1,683	520	78.42	26,015	4.47
C260816	106,140	106,707	577	5	90.55	715	80.14
scaffold2141	106,428	106,676	4,122	3,874	85.54	26,015	0.96
C247364	106,760	107,181	425	2	93.16	452	93.81
C154602	107,116	107,250	1	131	96.3	131	100
C161564	107,116	107,250	1	135	97.78	135	100
C257862	107,185	107,834	633	1	90.18	634	99.84
scaffold2141	107,443	107,834	392	1	92.62	26,015	1.51
C236392	107,768	108,106	348	10	95.87	348	97.41
C148258	108,200	108,324	1	125	97.6	125	100
C208876	108,200	108,429	1	230	96.52	230	100
C180808	108,263	108,429	6	172	98.8	172	97.09
C180810	108,263	108,429	6	172	98.2	172	97.09
scaffold753	108,385	110,620	2,662	372	92.39	2,662	86.06
C251496	108,399	108,911	511	1	91.42	511	100
C260740	108,399	109,115	1	713	90.17	713	100
C253776	108,763	108,981	480	262	91.82	548	39.96
scaffold2141	108,878	109,186	392	84	91.94	26,015	1.19
C188894	109,069	109,255	1	187	93.58	187	100
C188896	109,069	109,255	1	187	95.19	187	100
C253776	109,099	109,293	261	66	96.94	548	35.77
C218542	109,271	109,535	259	1	92.08	259	100
C190164	109,290	109,482	4	190	91.19	190	98.42
C215364	109,416	109,664	1	249	97.19	249	100
C211276	109,604	109,840	1	237	98.31	237	100
C181396	109,774	109,947	174	1	100	174	100
C189544	109,881	110,070	1	189	95.26	189	100
C256080	109,982	110,199	351	558	86.7	593	35.08
scaffold1065	110,001	110,199	3,977	3,788	87	6,582	2.89
C171802	110,003	110,158	155	1	92.95	155	100
C213902	110,092	110,335	244	1	96.72	244	100

C151573	110,269	110,394	128	1	97.66	128	100
C175412	110,335	110,489	1	157	93.63	163	96.32
C213064	110,335	110,570	1	238	92.44	242	98.35
C150739	110,349	110,476	1	128	99.22	128	100
C211716	110,413	110,642	4	233	95.65	238	96.64
C167642	110,429	110,570	146	5	96.48	146	97.26
scaffold1642	110,498	112,193	388	2,028	91.71	3,278	50.06
C191054	110,519	110,710	192	1	91.19	192	100
C154794	110,581	110,712	132	1	97.73	132	100
scaffold753	110,646	110,759	182	69	98.25	2,662	4.28
C181230	110,685	110,857	173	1	99.42	173	100
C159508	110,791	110,925	135	1	97.04	135	100
C228500	110,791	111,094	301	1	93.09	301	100
C146168	110,878	110,999	1	123	96.75	123	100
C145751	110,879	110,998	122	1	95.9	122	100
C146516	110,879	110,998	123	1	95.12	123	100
C172380	110,934	111,082	155	7	93.29	156	95.51
C166436	111,037	111,170	10	143	97.76	143	93.71
C137457	111,104	111,213	1	110	97.27	110	100
C159840	111,158	111,292	135	1	98.52	135	100
C149030	111,167	111,292	126	1	96.03	126	100
C257750	111,233	111,863	1	631	97.15	632	99.84
C258050	111,452	111,905	1	445	91.19	639	69.64
C147372	111,798	111,921	1	124	95.97	124	100
C156366	111,855	111,988	134	1	94.78	134	100
C157826	111,923	112,051	2	135	95.52	135	99.26
C188342	111,990	112,175	186	1	99.46	186	100
C231168	111,990	112,304	315	1	93.67	315	100
scaffold1682	112,025	113,408	6,776	5,394	93.86	9,838	14.06
scaffold1682	112,109	114,423	9,838	7,520	95.91	9,838	23.57
C168740	112,238	112,385	1	148	98.65	148	100
C180698	112,326	112,497	172	1	98.84	172	100
C208864	112,326	112,552	230	4	96.04	230	98.7
C193184	112,820	113,015	1	196	96.94	196	100
C209302	112,820	113,050	1	231	97.84	231	100
scaffold1682	112,937	113,115	7,385	7,207	96.65	9,838	1.82
C129766	112,995	113,095	101	1	99.01	101	100
C256080	113,017	113,265	1	249	93.17	593	41.99
scaffold1682	113,576	113,766	5,228	5,038	80.83	9,838	1.94

C242232	114,033	114,423	2	393	94.64	397	98.74
scaffold1682	115,168	116,598	2,296	884	90.26	9,838	14.36
C277238	115,961	116,131	630	460	94.15	1,997	8.56
C276834	116,422	117,366	1,056	66	81.53	1,921	51.59
scaffold1682	117,361	118,095	269	1,011	89.11	9,838	7.55
scaffold1330	118,227	119,029	40,823	40,021	99.75	50,208	1.6
C276834	119,034	119,851	1,127	1,921	89.54	1,921	41.38
C257424	121,379	121,512	131	1	92.59	624	20.99
C142195	121,395	121,512	117	1	96.64	117	100
C223164	121,458	121,734	1	277	98.92	277	100
C219322	121,467	121,728	262	1	94.27	262	100
C197128	121,668	121,866	204	1	94.61	204	100
C234384	121,857	122,172	335	21	92.77	335	94.03
C203748	122,125	122,317	194	1	94.39	218	88.99
C222480	122,125	122,375	250	1	92.91	274	91.24
C244938	122,377	122,798	1	422	93.36	425	99.29
C193300	122,504	122,699	196	1	95.92	196	100
C170308	122,633	122,784	152	1	97.37	152	100
C178614	122,633	122,798	169	4	96.99	169	98.22
C146390	122,965	123,087	123	1	99.19	123	100
C251118	123,024	123,516	1	504	90.91	504	100
C264294	123,046	123,881	835	1	90.15	835	100
C179324	124,021	124,190	170	1	93.53	170	100
C189580	124,021	124,206	189	4	94.09	189	98.41
C244938	124,129	124,548	422	3	94.76	425	98.82
scaffold2941	125,058	126,130	6,536	5,460	95.27	7,256	14.84
scaffold2941	125,771	127,798	5,517	3,503	93.8	7,256	27.77
scaffold2941	126,474	126,885	2,175	1,764	94.66	7,256	5.68
scaffold2941	127,946	128,308	3,392	3,012	91.1	7,256	5.25
C161948	128,178	128,308	1	130	97.71	135	96.3
scaffold2941	128,314	129,359	2,082	1,036	93.9	7,256	14.43
scaffold2941	129,559	130,220	1,127	481	92.6	7,256	8.92

Chapter 4. CRYPTIC ASEXUAL REPRODUCTION IN CAENORHABDITIS NEMATODES REVEALED BY INTERSPECIES HYBRIDIZATION[2]

4.1 INTRODUCTION

Most animal species reproduce by sex. Theory predicts there are advantages to being able to switch reproduction between sexual and asexual modes. However, facultative sex is rarely observed in animals, implying that there are strong selective pressures that prevent asexuality arising from an obligately sexual ancestor. One of the critical steps in the evolution of asexuality from a sexual ancestor is the transition from haploid to diploid maternal inheritance. Here we report that interspecific hybridization between two sexual *Caenorhabditis* nematode species (*C. nouraguensis* females and *C. becei* males) results in two classes of viable offspring. The first class consists of fertile offspring, which are produced asexually by sperm-dependent parthenogenesis (also called gynogenesis or pseudogamy); these progeny inherit a diploid maternal genome but fail to inherit a paternal genome. The second class consists of sterile hybrid offspring, which inherit both a diploid maternal genome and a haploid paternal genome. Using whole-genome sequencing of individual viable worms, we show that diploid maternal inheritance in both asexually produced and hybrid offspring results from the inheritance of two randomly selected homologous chromatids from *C. nouraguensis* oocytes. This genetic mechanism of diploid maternal inheritance is indistinguishable from that of many obligately asexual species. Furthermore, we show that intraspecies *C. nouraguensis* crosses can also result

[2] This chapter is closely adapted from (Lamelza *et al.* 2019), which was done in close collaboration with Janet Young in Harmit Malik's lab at the Fred Hutchinson Cancer Research Center (*C. nouraguensis* assembly and bioinformatics) and Luke Noble in Matthew Rockman's lab at NYU (*C. becei* assembly).

in a low frequency of asexual reproduction through diploid maternal inheritance. Thus, *C. nouraguensis* provides a genetically tractable model to study the evolutionary origins of asexuality from obligately sexual species.

4.2 RESULTS

4.2.1 *Reciprocal C. nouraguensis x C. becei crosses exhibit distinct F1 embryonic lethal phenotypes*

C. nouraguensis and *C. becei* are obligately outcrossing sister species consisting of female and male individuals (Figure 4.1A). We initially studied the hybridization of *C. nouraguensis* strain JU1825 with *C. becei* strain QG711. The mating of individuals within the same strain results in a high proportion of viable progeny (Figure 4.8, Videos S1 and S2). However, when we mated *C. becei* females (QG711) to *C. nouraguensis* males (JU1825), we found that all F1 embryos die during development (Figure 4.1B). DIC time-lapse imaging shows that these F1 embryos stereotypically arrest and fail to undergo further divisions at approximately the 32-cell stage (Video S3). In the reciprocal cross in which we mated *C. nouraguensis* females (JU1825) to *C. becei* males (QG711), we found that most F1 embryos also die during development, but the dead F1 embryos have a high variance in the range of phenotypes and developmental stages when cell division arrests. Some embryos arrest early in development (approximately 1-4 cell stage) whereas some arrest at later stages (approximately 44-87 cell stage) (Video S4).

4.2.2 *Crossing C. nouraguensis females to C. becei males results in rare viable F1 progeny*

Interestingly, a small fraction of F1 progeny from *C. nouraguensis* female x *C. becei* male crosses develop into viable adults. Such viable progeny were found using three different *C.*

nouraguensis strains and two *C. becei* strains (Figure 4.9), indicating that the production of viable progeny is a general feature of *C. nouraguensis* female x *C. becei* male crosses. By contrast, no surviving F1 were found in any of the reciprocal *C. becei* female x *C. nouraguensis* male crosses (Figure 4.1B and Figure 4.9). Additionally, crosses between *C. nouraguensis* females and males of other closely related *Caenorhabditis* species gave no surviving F1 progeny (Figure 4.9), indicating that the production of viable F1 progeny is specific to the *C. nouraguensis* female x *C. becei* male cross. We tested the fertility of each viable F1 animal by mating to a *C. nouraguensis* individual of the opposite sex and determining whether they produced F2 embryos (Figure 4.1C). We found that ~2/3 of F1 tested are sterile, but the other ~1/3 were fertile (JU1825 x QG711: 18 sterile, 8 fertile; NIC54 x QG711: 27 sterile, 10 fertile) (Figure 4.1D).

After determining their fertility, we tested if the rare viable F1 were indeed interspecies hybrids at a genetic level. We genotyped each F1 at one of two autosomal loci using PCR-restriction digest assays that distinguish between *C. nouraguensis* and *C. becei* alleles (Figure 4.1C). We found that about two-thirds of viable F1 inherited both parental alleles and therefore are genetic hybrids (JU1825 x QG711: 17/26 F1; NIC54 x QG711: 26/37 F1). However, to our surprise, we found that the remaining viable F1 adults carry only the maternal *C. nouraguensis* allele and no paternal *C. becei* allele (JU1825 x QG711: 9/26 F1; NIC54 x QG711: 11/37 F1) (Figure 4.1D). Interestingly, the absence of the paternal DNA correlates with fertility: all fertile F1 had only the maternal *C. nouraguensis* allele and not the paternal *C. becei* allele, while all but two sterile F1 had a hybrid genotype, with alleles from both parents. We observed a similar correlation between genotype and fertility among rare F1 when a different paternal *C. becei* strain was used (Figure 4.10).

The presence of sterile F1 hybrids in these crosses is not especially surprising, as many interspecies hybridizations result in sterile F1 hybrids (Coyne and Orr 2004). However, it is surprising to find fertile F1 that appear to have inherited only maternal *C. nouraguensis* DNA. These fertile F1 behave like *C. nouraguensis* when backcrossed to either parental species. Backcrossing fertile F1 to *C. nouraguensis* results in many viable F2 progeny that develop to adulthood. By contrast, backcrossing fertile F1 to *C. becei* results in almost all F2 dying during embryogenesis other than the few rare viable F2 progeny in crosses to *C. becei* males (as in the original interspecies cross). Thus, it appears that interspecies hybridization between *C. nouraguensis* females and *C. becei* males produces a low frequency of fertile F1 that carry only maternal *C. nouraguensis* DNA.

The production of fertile F1 progeny with a maternal genotype made us wonder whether *C. nouraguensis* females might generate sperm and reproduce as hermaphrodites at a low frequency. If this were the case, virgin females should also be able to produce progeny. We monitored 60 virgin females from two *C. nouraguensis* strains (JU1825 and NIC59) and two *C. becei* strains (QG704 and QG711) for five days but failed to find any progeny, indicating that production of fertile F1 is not due to cryptic hermaphroditism, but requires mating to *C. becei* males. We hypothesize that fertile F1 are the result of gynogenetic reproduction, meaning that *C. nouraguensis* oocytes require fertilization by *C. becei* sperm to initiate development, but do not inherit the paternal *C. becei* genome.

4.2.3 *Fertile interspecific F1 females are diploid*

How do the fertile F1 compensate for the lack of a paternal haploid genome? We presumed that the autosomes of the fertile F1 had to be at least diploid to be viable (Sadler and

Shakes 2000), because all well-studied *Caenorhabditis* species have five diploid autosomes, in addition to an X chromosome that is either diploid (females) or haploid (males) (Cutter *et al.* 2009; Fierst *et al.* 2015; Roelens *et al.* 2015).

We determined the ploidy of F1 females by counting the number of chromosomes in their oocytes. In *C. elegans*, the female germline produces oocytes containing chromosomes that have undergone prophase of meiosis I (DNA replication and recombination) but have yet to undergo the two meiotic divisions. The oocyte contains six bivalents, each of which is composed of a pair of homologous chromosomes and their sister chromatids held together by a combination of recombination and sister chromatid cohesion (Figure 4.2A). Therefore, diploid individuals have six bivalents in their oocytes, which appear as six distinct DAPI-staining bodies (Figure 4.2B). *C. elegans* exhibits complete crossover interference, meaning that meiotic recombination occurs only once per homologous set of chromosomes (Meneely *et al.* 2002). Thus, a triploid individual will have two homologous chromosomes that recombine and form a bivalent as well as a non-recombinant homolog as a univalent. Therefore, triploid individuals have six bivalents and six univalents in mature oocytes, which appear as twelve DAPI-staining bodies (Figure 4.2B) (Madl and Herman 1979).

We crossed *C. nouraguensis* (JU1825) females to *C. becei* (QG711) males, collected fertile F1 females and dissected and stained their germlines with DAPI. As controls, we examined the germlines of JU1825 and QG711 females after mating with conspecific males. Both parental species mostly have six DAPI-staining bodies, consistent with six diploid chromosomes (like *C. elegans*). Most fertile F1 females also have six DAPI-staining bodies, indicating they are diploid (Figure 4.2C).

4.2.4 *Fertile interspecific F1 inherit two random homologous chromatids from each maternal C. nouraguensis bivalent*

There are three mechanisms by which the fertile F1 animals could have diploid maternal inheritance. They could inherit two chromatids from each maternal bivalent in a *C. nouraguensis* oocyte, inherit one chromatid from each bivalent and undergo genome-wide diploidization by endoreplication, or inherit the original two maternal chromatids by producing diploid eggs by mitosis (apomixis). To distinguish between these possibilities, we used whole-genome sequencing of individual worms to determine the genetic identity of the chromatids inherited by fertile F1 individuals.

All four of the chromatids within a bivalent are genetically distinct; of the six possible ways of combining two chromatids, four have distinct chromosomal genotypes and two are indistinguishable by Illumina whole-genome sequencing (Figure 4.3A, 4.3B and Figure 4.12). Thus, the genotype of the fertile F1 can indicate which two chromatids were inherited. Because recombination in *Caenorhabditis* species occurs predominantly on chromosome arms (Rockman and Kruglyak 2009; Ross *et al.* 2011), inheriting two sister chromatids would usually result in homozygosity in the centers of chromosomes and heterozygosity on one of the arms, whereas inheriting two homologous chromatids would result in heterozygosity in the center of each chromosome (Figure 4.3B). By contrast, endoreplication of single chromatids would lead to homozygosity along all chromosomes, and apomixis would lead to heterozygosity along all chromosomes.

We first crossed two genetically distinct *C. nouraguensis* strains, NIC59 and JU1825, to generate heterozygous *C. nouraguensis* (NIC59/JU1825) females, which we then crossed to *C. becei* (QG711) males. We collected eleven fertile F1 and performed whole-genome amplification and sequencing of each one (Figure 4.3A). We determined the genotype of each fertile

individual's chromosomes by calculating average NIC59 SNP frequencies in 50-kb windows across their genome (see Materials and Methods for details). An approximately 0.50 NIC59 allele frequency was inferred as heterozygosity (NIC59/JU1825), whereas an approximately 1.0 or 0.0 NIC59 allele frequency was inferred as homozygous NIC59 or JU1825, respectively. This experiment also allowed us to determine whether fertile F1 inherit only maternal *C. nouraguensis* DNA or whether some paternal *C. becei* DNA is also present.

We found that nearly 100% of each fertile F1's reads map to the *C. nouraguensis* genome assembly; only a small percentage (0.1-0.2%) map to the *C. becei* genome assembly (Table 4.8 and Figure 4.13). These results resemble our controls that contain only *C. nouraguensis* DNA (samples: F1_NIC59_JU1825 and NIC59plusJU1825, Table 4.1 and Figure 4.13). Therefore, our sequencing results confirm that fertile F1 inherit only maternal *C. nouraguensis* DNA. Strikingly, whole-chromosome genotyping data shows that the fertile F1 always inherited two homologous chromatids from each maternal bivalent (Figure 4.3C, Figure 4.12 and Figure 4.13). Combining chromosome genotypes from the fertile F1 indicates that two randomly selected homologous chromatids were inherited (Figure 4.3D). Based on these findings, we can rule out endoreplication and apomixis. Instead, we conclude that diploid, fertile F1 inherit two homologous chromatids from each maternal bivalent in *C. nouraguensis* oocytes.

4.2.5 *Sterile F1 inherit a diploid C. nouraguensis genome and a haploid C. becei genome*

Having established that fertile F1 adults inherit only a diploid maternal genome, we next examined the genome-wide inheritance pattern in sterile F1 adults. We performed whole-genome sequencing of ten sterile F1 adults derived from the same interspecies crosses mentioned in the previous section (Figure 4.3A).

Surprisingly, five of the sterile F1 appear to be triploid hybrids that inherited a diploid *C. nouraguensis* maternal genome and a haploid *C. becei* paternal genome (females F1_6, F1_12, F1_20; males F1_17 and F1_21, Table 4.1). In these individuals, the normalized coverage of *C. nouraguensis* autosomes was approximately double that of the *C. becei* autosomes (Figure 4.4A). Furthermore, genotyping of the *C. nouraguensis* chromosomes revealed that these individuals inherited two homologous chromatids from each maternal *C. nouraguensis* bivalent, just like the fertile F1 (Figure 4.4B and Figure 4.13). Three other sterile F1 males are also hybrids that inherited two homologous chromatids from each maternal bivalent. However, they are unlikely to be true, full triploids because the normalized coverage of the *C. becei* autosomes is considerably less than half that of the *C. nouraguensis* autosomes (males: F1_10, F1_16 and F1_23, Figure 4.4A). Instead, we hypothesize that the reduced coverage of the *C. becei* genome results from mosaicism, in which only a subset of cells in these individuals contain the paternal *C. becei* DNA. The *C. becei* autosomes within each individual hybrid mosaic exhibit similar levels of coverage (except perhaps Chr. V of F1_16), suggesting that most hybrid cells within these individuals have a complete haploid copy of the *C. becei* autosomal genome (diploid-triploid mosaic hybrids) (Figure 4.4A).

Combining maternal chromosome genotypes from the sterile F1 indicates that two randomly selected homologous chromatids were inherited (Figure 4.4C and Figure 4.12C). Additionally, the combined data on maternal genotype frequencies from both sterile and fertile F1 also indicate random homologous chromatid inheritance (Figure 4.4D and Figure 4.12C). These results show that both the fertile and sterile F1 share the same mechanism of diploid maternal inheritance and are distinguished by whether they inherited the haploid *C. becei* genome.

4.2.6 *The C. becei X-chromosome is toxic to F1 hybrids*

In *Caenorhabditis*, males are hemizygous for the X-chromosome (XO) and produce haploid sperm that carry either a single X or no X. Therefore, roughly half of the F1 are expected to inherit the paternal *C. becei* X-chromosome. However, despite inheriting *C. becei* autosomes, none of the sterile F1 we sequenced show inheritance of the *C. becei* X-chromosome (Figure 4.13). We collected additional viable F1 hybrids from the interspecies cross and determined whether any inherited the *C. becei* X-chromosome by PCR genotyping of two indel polymorphisms, one X-linked and one autosomal, that distinguish between *C. nouraguensis* and *C. becei* sequences (Figure 4.5A). We found that none of the F1 inherited the *C. becei* X-chromosome (Figure 4.5B), regardless of whether they were autosomal hybrids (22 females and 12 males) or had a maternal genotype (10 females).

Our failure to find any viable F1 carrying a *C. becei* X-chromosome suggests that it is toxic to F1 hybrids. Alternatively, this pattern could reflect an unusual segregation pattern specifically involving the X chromosome in hybrids. To distinguish between these possibilities, we PCR genotyped dead F1 embryos from the same interspecies cross. Consistent with the expected inheritance of the X-chromosome from males, we found that 50% of dead F1 embryos inherited both the *C. becei* paternal X-chromosome and the *C. nouraguensis* maternal X-chromosome (17/34), whereas 47% inherited only the maternal *C. nouraguensis* X-chromosome (16/34). One dead F1 embryo possessed only a paternal *C. becei* genotype, suggesting that it lost the maternal *C. nouraguensis* X-chromosome. Because dead F1 embryos can inherit the *C. becei* X-chromosome but viable F1 animals do not, we conclude that the *C. becei* X-chromosome is toxic to F1 hybrids. Previous work has described the “large X-effect” on hybrid inviability and

sterility (Coyne and Orr 2004). We hypothesize that a dominant incompatibility involving one or more loci on the *C. becei* X chromosome underlies its toxic effect on hybrids with *C. nouraguensis*.

4.2.7 *Dead F1 embryos have unusual maternal and paternal inheritance*

To determine whether the unusual patterns of maternal and paternal inheritance seen in viable F1 also occur in the dead F1 embryos in the interspecies cross, we PCR genotyped the same dead F1 embryos described above at an autosomal indel polymorphism that distinguishes the two species (Figure 4.5C). We found that 27 of the 35 dead F1 embryos (77%) had a heterozygous *C. nouraguensis*/*C. becei* genotype while eight (22%) had an only maternal *C. nouraguensis* genotype (Figure 4.5C). These data suggest that although most embryos inherit both maternal and paternal DNA, a significant fraction fail to inherit paternal *C. becei* DNA for this marker. Aneuploidy caused by incomplete loss of the paternal genome may contribute to death of embryos.

To determine whether dead F1 embryos also inherit two homologous *C. nouraguensis* chromatids from each maternal bivalent, we PCR genotyped them at an autosomal indel polymorphism in the center of chromosome I that distinguishes NIC59 and JU1825 alleles. We found that 34/36 dead F1 embryos had a heterozygous NIC59/JU1825 genotype, whereas one embryo had a JU1825 genotype and one embryo had a NIC59 genotype (Figure 4.5D). These data suggest that very few embryos are derived from a canonical female meiosis in which a haploid maternal complement is inherited. Instead, most embryos appear to have inherited at least two homologous chromatids from an aberrant female meiosis.

To better understand how female meiosis is modified so that at least two homologous chromatids are inherited, we quantified the number of polar bodies found in F1 embryos, of which the vast majority are expected to die. If most *C. nouraguensis* oocytes fertilized by *C. becei* sperm have aberrant meiotic segregations that result in the inheritance of two homologous chromatids, we might expect abnormalities in polar body number or morphology. We DAPI stained F1 embryos derived from crossing *C. nouraguensis* (JU1825) females to *C. becei* (QG711) males. As a control, we stained embryos derived from intraspecies JU1825 crosses. As expected, JU1825 embryos always have two polar bodies whose morphology and size appear similar across individuals (Figure 4.5E and 4.5I). By contrast, hybrid embryos often have fewer than two polar bodies, potentially indicating failed or abnormal meiotic segregations. The polar bodies in hybrid embryos also often appear abnormal in their morphology (Figure 4.5F-I). Together with our genotyping data, these observations indicate that meiotic segregations are frequently perturbed in *C. nouraguensis* females mated to *C. becei* males.

4.2.8 *Diploid maternal inheritance occurs occasionally in C. nouraguensis intraspecies crosses*

To determine whether diploid inheritance from *C. nouraguensis* oocytes can occur independently of interspecies hybridization, we designed an experiment in which *C. nouraguensis* sperm can fertilize *C. nouraguensis* oocytes but cannot contribute a paternal genome. In this scenario, a canonical female meiosis would result in haploid maternal genome inheritance and the embryos would fail to develop into viable adults (Sadler and Shakes 2000). However, if *C. nouraguensis* oocytes inherit a diploid maternal genome, they could complete embryogenesis and develop into viable adults.

We blocked the inheritance of the haploid sperm genome using UV irradiation to damage and inactivate the sperm genome while still allowing for fertilization and development of the maternal haploid embryo (Goda *et al.* 2006; Fopp-Bayat *et al.* 2017). We exposed NIC59 and JU1825 adult males to a high dose of shortwave UV radiation and then crossed them to unexposed JU1825 or NIC59 adult virgin females, respectively. We then screened for and PCR genotyped viable progeny at several autosomal loci to determine whether they inherited both maternal and paternal DNA (heterozygous NIC59/JU1825), or just maternal DNA (Figure 4.6A).

First, we crossed JU1825 females to NIC59 UV-irradiated males. Only seven out of approximately 17,800 embryos survived to adulthood. Upon PCR genotyping a chromosome I indel polymorphism, we found that four of these seven individuals had a heterozygous NIC59/JU1825 genotype and therefore inherited both maternal and paternal alleles at this locus. We presume these animals represent rare cases in which the sperm genome was not destroyed. However, three females inherited only the maternal JU1825 allele (Figure 4.6B). We PCR genotyped these females at two other unlinked autosomal indel polymorphisms. All had a maternal JU1825 genotype at all three loci, suggesting that they inherited only the maternal genome (Figure 6C). We performed similar experiments with NIC59 females and UV-irradiated JU1825 males. Only two out of roughly 7,120 progeny survived to adulthood. These two females had only maternal NIC59 genotypes at all three autosomal loci (Figure 4.6B and 4.6C). These results indicate that diploid maternal inheritance in *C. nouraguensis* oocytes does not require fertilization by *C. becei* sperm and can occur even in intraspecies matings.

4.3 Discussion

In this study we investigated the hybridization of two sexual *Caenorhabditis* species, *C. nouraguensis* and *C. becei*. We found that although most offspring of *C. nouraguensis* females

crossed to *C. becei* males die during embryogenesis, rare viable offspring are produced. About one-third of the viable offspring are fertile and result from a combination of diploid maternal inheritance and paternal genome loss, two traits that define gynogenetic reproduction. The remaining two-thirds of the viable offspring are sterile and possess a diploid maternal *C. nouraguensis* genome together with a haploid paternal *C. becei* genome. However, none of the sterile hybrids inherit the *C. becei* X-chromosome, although it is inherited at the expected frequency in dead embryos. Finally, we found that diploid maternal inheritance also occurs at a low frequency in intraspecies *C. nouraguensis* crosses. We hypothesize that the same mechanism of diploid maternal inheritance occurs in both intraspecies and interspecies crosses, but that fertilization of *C. nouraguensis* oocytes by *C. becei* sperm dramatically increases its frequency, suggesting that a signal from *C. becei* sperm alters *C. nouraguensis* female meiosis.

4.3.1 *Mechanism of diploid maternal inheritance*

We found that diploid maternal inheritance in fertile and sterile offspring almost always results from inheriting two randomly selected homologous chromatids from each maternal bivalent in the oocyte. We also found that dead F1 embryos almost always inherit at least two homologous chromatids from *C. nouraguensis* oocytes. Thus, diploid maternal inheritance is a general feature of *C. nouraguensis* oocytes when fertilized by *C. becei* sperm. We hypothesize that this diploid maternal inheritance reflects a stereotypical segregation mechanism that is a slight modification of canonical meiotic divisions (Figure 4.7A).

In our model (Figure 4.7B), meiotic prophase occurs normally in *C. nouraguensis* oocytes, generating six recombinant bivalents. Upon fertilization by *C. becei* sperm, the bivalents randomly bi-orient their homologs on the meiotic spindle, sister chromatid cohesion is lost

between homologous chromosomes, and homologs segregate at anaphase I. However, instead of one set of homologs getting segregated into the first polar body, both sets are retained in the oocyte. Both half-bivalents within the oocyte then randomly bi-orient on the meiotic spindle, sister chromatid cohesion is lost, and sister chromatids segregate at anaphase II. One chromatid from each half-bivalent segregates into a polar body while the other is retained in the oocyte. Thus, the oocyte inherits two randomly selected homologous chromatids from each bivalent. This model suggests that diploid maternal inheritance could result from a modification of female meiosis such as failure of cytokinesis after anaphase I. There is evidence that incomplete anaphase I and a failure of cytokinesis generates diploid eggs in *Daphnia pulex*, which undergoes obligate pathenogenesis (Hiruta *et al.* 2010).

Some loss-of-function mutations and gene knockdowns in *C. elegans* can possibly lead to diploid maternal inheritance (Severson *et al.* 2009; McNally *et al.* 2016), suggesting plausible cellular mechanisms underlying this phenomenon. During normal meiosis I in *C. elegans*, the meiotic spindle migrates and orients perpendicularly to the oocyte cortex. In anaphase I, one set of homologs segregates towards the cortex into the first polar body while the other is retained in the oocyte. However, weak knockdown of dynein heavy chain (*dhc-1*) causes the meiotic spindle to be oriented parallel to the cortex rather than perpendicular during meiosis I; as a result, homologous chromosomes segregate parallel to the cortex in anaphase I and no polar body is formed (McNally *et al.* 2016). The half-bivalents that remain in the oocyte then segregate relatively normally during anaphase II and form a polar body. The oocyte presumably inherits two randomly selected homologous chromatids from each bivalent. A similar misalignment of the meiotic spindle at metaphase I is also thought to be the mechanism underlying the generation of diploid eggs in the obligately parthenogenetic species *Drosophila mangabeirai* (Murdy and

Carson 1959). We hypothesize that a misalignment of the meiotic spindle or a failure of cytokinesis may similarly result in diploid maternal inheritance in *C. nouraguensis*.

4.3.2 *Paternal genome loss*

In addition to diploid maternal inheritance, gynogenesis also requires paternal genome loss. In viable F1 offspring from the *C. nouraguensis* x *C. becei* cross, we found that the entire haploid *C. becei* paternal genome can be inherited in all cells (triploid hybrids), only a subset of cells (diploid-triploid mosaic hybrids), or in no cells (fertile gynogenetic offspring). Interestingly, the offspring of naturally gynogenetic fish species can have a similar spectrum of paternal inheritance. Specifically, most offspring completely fail to inherit the paternal genome (diploid maternal clones), but a small fraction of offspring can inherit it in a subset (diploid-triploid mosaic) or all cells (triploid) (Goddard *et al.* 1998; Lamatsch *et al.* 2002; Itono *et al.* 2007). Paternal genome loss can occur in a variety of contexts and has different underlying causes, each of which could potentially explain the loss of the *C. becei* genome (Baker 1975; Nelson-Rees *et al.* 1980; Yasuda *et al.* 1995; Werren *et al.* 2008; Yamaki *et al.* 2016; Aldrich *et al.* 2017). In the context of interspecies hybridization, detailed studies in *Hordeum* (barley) and *Xenopus* show that paternal chromosomes fail to recruit centromeric proteins and lag during embryonic divisions (Sanei *et al.* 2011; Gibeaux *et al.* 2018).

4.3.3 *Characteristics of incipient gynogenesis*

Asexuality is rare in animals, suggesting that there are strong selective pressures that prevent its evolution from a sexual ancestor. However, it is largely unknown how the transition from sexuality to asexuality might occur, nor how the costs and benefits of such a transition

might influence the longevity of a nascent asexual lineage. Some obligately gynogenetic species are composed entirely of females, and require that their eggs be fertilized and activated by males of a closely related species (Beukeboom and Vrijenhoek 1998). We hypothesize that the rare gynogenetic reproduction seen in *C. nouraguensis* x *C. becei* hybridizations could serve as a transitional state between sexual and obligate or facultative interspecies gynogenetic reproduction. Furthermore, we hypothesize that the diploid maternal inheritance observed in *C. nouraguensis* intraspecies crosses could be used as a stepping-stone towards either gynogenetic or parthenogenetic reproduction. Both interspecies gynogenesis and intraspecies diploid maternal inheritance could be features of a nascent asexual lineage.

We have shown that diploid maternal inheritance in *C. nouraguensis* oocytes is the result of inheriting two homologous chromatids that have undergone meiotic recombination, and therefore fits a model of automixis in which two homologous chromatids are combined ("central fusion"). Because there is only a single recombination event that is biased towards the chromosome ends, inheriting two homologous chromatids either maintains heterozygosity across the entire chromosome or most of it (Figure 4.3B). However, recombination does occur in the middle of chromosomes at a lower frequency (Rockman and Kruglyak 2009), which will eventually result in genome-wide homozygosity and inbreeding depression after many generations of automixis. Most *Caenorhabditis* species appear to be dioecious (females and males), with some exhibiting genetic hyperdiversity (Cutter *et al.* 2006; Dey *et al.* 2013). Therefore, inbreeding depression could be a sizable hurdle to overcome in the evolution of this form of gynogenesis from an obligately sexual *Caenorhabditis* ancestor.

Given that two copies of each maternal autosome are almost always inherited in the interspecies *C. nouraguensis* x *C. becei* cross, we were surprised to observe that some diploid

fertile F1 are males. Indeed, sequencing data show that the three fertile males have only a single X chromosome (Figure 4.12 and Figure 4.13). Thus, missegregation of the X-chromosome during the modified female meiosis can result in only one X chromatid being inherited by the oocyte. Interestingly, two of the three fertile males were also aneuploid (triploid) for a single *C. nouraguensis* autosome, indicating that widespread chromosome missegregation occurred while producing these individuals. Generating males at a low frequency can theoretically aid the spread and maintenance of gynogenesis by transmitting alleles required for diploid maternal inheritance into neighboring sexual populations (Jaquiéry *et al.* 2014). The production of males even at a low frequency can also help propagate an intraspecific gynogenetic lineage by enabling the activation of the egg and inheritance of centrioles without relying on males from a related species, as seen in *Mesorhabditis belari* (Grosmaire *et al.* 2019). Thus, it is notable that gynogenetic reproduction in *C. nouraguensis* x *C. becei* hybridizations already exhibits the ability to produce males.

In conclusion, our study establishes *C. nouraguensis* as a new genetic model system to study rare asexuality in animals. Further study of this system may provide insights into the mechanisms of diploid maternal inheritance and paternal genome loss, as well as the obstacles associated with a transition from sexual to asexual reproduction.

4.4 MATERIALS AND METHODS

4.4.1 *Strain isolation and maintenance*

All strains of *Caenorhabditis* used in this study were derived from single gravid females isolated from rotten fruit or flowers (Kiontke *et al.* 2011; Stevens *et al.* 2018). Strains of *C. nouraguensis* were kindly provided by Marie-Anne Félix (“JU” prefix) and Christian Braendle

("NIC" prefix). Most strains are outbred, except JU2079, QG2082 and QG2083, which are inbred lines derived from JU1825, QG704 and QG711, respectively. These inbred lines were used to generate the *C. nouraguensis* and *C. becei* genome assemblies. Strain stocks were stored at -80°. Strains were maintained at 25° on standard NGM plates spread with a thin lawn of *E. coli* OP50 bacteria (Brenner 1974).

4.4.2 *Phylogenetic analysis*

To construct a phylogeny, we used a subset of the coding sequences from Kiontke et al. (Kiontke *et al.* 2011). The following genes were used: *ama-1, lin-44, par-6*, *pkc-3*, *W02B12.9*, *Y45G12B.2a*, *ZK686.3*, and *ZK795.3*. The *C. yunquensis* coding sequences were obtained from NCBI, while all other coding sequences were obtained from the caenorhabditis.org BLAST server (blast.caenorhabditis.org/#) using the *C. elegans* homolog as the query sequence. The coding sequences for each species were concatenated and aligned with MUSCLE (Edgar 2004) using default settings. A maximum likelihood phylogeny was made using PhyML 3.0 online (http://www/atgc-montpellier.fr/phyml/) (Guindon *et al.* 2010) using a Generalized Time Reversible (GTR) substitution model with six substitution rate categories. Tree searching was performed using Nearest Neighbor Interchanges (NNIs). Branch support was determined by 100 bootstraps.

4.4.3 *Quantifying strain viability*

Strain viability was quantified as in (Lamelza and Ailion 2017). Briefly, we crossed 10 virgin L4 females and males of the same strain on a single plate coated with palmitic acid. The plates were placed at 25° overnight, allowing the worms to mature to adulthood and begin

mating. The adult worms were then moved to a new plate rimmed with palmitic acid, allowed to mate and lay eggs for 1.5 hours at 25°, and then removed. The eggs laid within that time were immediately counted. Two days later, we placed the plates at 4° for 1 hr and counted the number of healthy L4 larvae and young adults per plate. We defined the percentage of viable progeny as the total number of L4 larvae and young adults divided by the total number of eggs laid.

4.4.4 *DIC imaging of embryogenesis*

Twenty-four hours after the L4 stage, we picked 20-40 gravid females into a 30 µl pool of 1x Egg Buffer (25 mM HEPES pH 7.4, 118 mM NaCl, 48 mM KCl, 2 mM EDTA, 0.5 mM EGTA) on a glass coverslip and dissected out embryos by cutting the adults in half using two needles like scissors. The embryos were then transferred with a glass capillary tube into a fresh pool of Egg Buffer to dilute any contaminating bacteria and then placed onto a 2% agarose pad on a glass slide. The mounted embryos were put into a humid chamber for 20 minutes before adding a coverslip. The edges of the coverslips were sealed with petroleum jelly to prevent evaporation of the agarose pad. DIC z-stack images were captured every 3 minutes for 10 hours using a Nikon Eclipse 80i compound microscope (60x oil lens, 1.40 NA). Images were processed using Image J (Collins 2007).

4.4.5 *Calculating the frequency of rare interspecific F1*

For each interspecies cross we set up 20-25 plates, each with one virgin L4 female crossed to a single virgin L4 male. The plates were placed at 25° overnight, allowing the worms to mature to adulthood and begin mating and laying eggs. To prevent overcrowding, the adult couples were moved to a new plate every day for three days. The number of eggs laid on each

plate was counted the day the parents were removed. Each plate was monitored for two days to check for the presence of rare viable F1, which were counted and moved onto a new plate.

4.4.6 *Generating rare interspecific F1 and testing their fertility*

To collect enough rare viable F1 animals for fertility testing and genotyping, we set up six cross plates, each with 30 *C. nouraguensis* L4 females and 30 *C. becei* L4 males and let them develop into adults overnight at 25°. We monitored the plates for the presence of viable F1 larvae each day for 5-6 days. Each viable F1 was moved to its own new palmitic acid-rimmed plate, and its fertility was tested as either an L4 or young adult by backcrossing it to one or two *C. nouraguensis* individuals of the opposite sex. F1 individuals were classified as sterile if they mated to *C. nouraguensis* but produced no F2 embryos, while F1 individuals were classified as fertile if they produced F2 embryos. Mating was inferred by the presence of a male-deposited copulatory plug on the female vulva (Palopoli *et al.* 2008). After fertility testing and PCR genotyping, we found that hybrid males produced no embryos when crossed to *C. nouraguensis* females and were classified as sterile. However, they often failed to produce a copulatory plug, so we cannot tell whether sterility is due to abnormal germline development or defective mating.

To track the fate of the two maternal homologous chromosomes in subsequent crosses, we crossed two genetically distinct strains of *C. nouraguensis* (JU1825 and NIC59) to generate heterozygous *C. nouraguensis* NIC59/JU1825 females. Specifically, we crossed virgin L4 JU1825 females to NIC59 males and vice versa to generate female offspring that are heterozygous NIC59/JU1825 at all nuclear loci and carry either JU1825 or NIC59 mitochondria, respectively. We denote the genotype of these heterozygous females by the following nomenclature: “(mitochondrial genotype); nuclear genotype”. Therefore, the first cross produces

C. nouraguensis (J); N/J females, while the second produces *C. nouraguensis* (N); N/J females (Table S1). We then crossed both types of *C. nouraguensis* N/J virgin females to *C. becei* QG711 males. We collected the rare viable F1 from these interspecies crosses and tested their fertility by backcrossing them to the *C. nouraguensis* strain matching their mitochondrial genotype in order to avoid known cytoplasmic-nuclear incompatibilities within this species (Lamelza and Ailion 2017).

4.4.7 *Generating worm and embryo lysates for PCR*

Single worm PCR: Each adult worm was placed in a PCR tube with 10 µl of lysis buffer (1x Phusion HF buffer + 1.0% Proteinase K) and frozen at -80° for at least 15 minutes. The worms were then lysed by incubating the tubes at 60° for 60 mins and 95° for 15 mins. 1 µl of worm lysate was used in each 10 µl PCR reaction. Control lysates were generated by mixing 10-15 worms in lysis buffer.

Single embryo PCR: Dead embryos were picked into a 30 µl pool of 2 mg/ml chitinase on an 18x18 mm coverslip. OP50 bacteria were washed off embryos by pipetting the chitinase solution up and down. Using a glass capillary tube, embryos were then transferred to a new pool of chitinase and washed again. Using a glass capillary tube, individual embryos were placed in a PCR tube with 5 µl of 2 mg/ml chitinase solution and incubated at room temperature for 2 hours. Afterwards, 5 µl of lysis buffer (2x Phusion HF Buffer + 2.0% Proteinase K) was added to each PCR tube and mixed by light vortexing. The samples were then frozen and lysed as in the single worm PCR protocol. 2.5 µl of embryo lysate was used in each 10 µl PCR reaction. Control lysates were generated by mixing 10-15 embryos in the same PCR tube. PCR primers are included in Table 4.2.

4.4.8 *Fixing and DAPI staining the female germline*

Fixing and DAPI staining germlines was performed as in (Bhalla and Dernburg 2005). F1 L4 females were mated to *C. nouraguensis* L4 males. 24 hours later, we dissected the gonads from fertile F1 females and stained with DAPI. 30-40 females were picked into 15 µl of 1x Egg Buffer + 0.1% Tween-20 on a glass coverslip and their heads were cut off with needles to expel the gonad. After dissection, 15 µl of Fixation Solution (1x Egg Buffer + 7.4% formaldehyde) was added and mixed well by pipetting. Gonads were fixed for 5 minutes, then 15 µl of the mixture was removed and a Histobond slide was laid down gently on top of the coverslip. Excess liquid was removed with Whatman paper and samples were frozen by placing the slide on a metal block on dry ice for 10 minutes. Coverslips were then quickly flipped off with a razor blade and the slides were placed in a Coplin jar with -20° methanol for 1 minute. The slides were then moved to a Coplin jar with 1x PBS + 0.1% Tween-20 for 10 minutes, then a Coplin jar with 1x PBS + 0.1% Tween-20 + 5 µl of 5 mg/ml DAPI for 15 minutes, then a Coplin jar with 1x PBS + 0.1% Tween-20 for 20 minutes. Samples were then mounted with 11 µl Mounting Media (35 µl 2M Tris + 15 µl MilliQ dH_2O + 450 µl glycerol) and sealed with nail polish. Images were taken using a Nikon Eclipse 80i compound microscope and processed using Image J.

4.4.9 *C. becei genome assembly and linkage map construction*

The chromosomal reference assembly for *C. becei* is based on high-coverage paired-end and mate-pair short-read sequencing, low-coverage Pacific Biosciences long reads, and genetic linkage information from an experimental intercross family. We used two inbred lines, QG2082 and QG2083, derived from isofemale lines QG704 and QG711 respectively by sib-mating for 25

generations. QG704 and QG711 were isolated from independent samples of rotten flowers (QG704) or fruit (QG711) collected in Barro Colorado Island, Panama, in 2012 (Stevens *et al.* 2018). We generated conservative *de novo* assemblies for the inbred lines, identified SNPs that distinguish them, and then used a six-generation breeding design to generate recombinant populations (G_4BC_2) from which we generated a SNP-based linkage map. With the map-scaffolded conservative assembly, we were able to evaluate more contiguous, less conservative genome assemblies for consistency, and we selected one of these as our final *C. becei* genome. Each of these steps is described in detail below.

We extracted genomic DNA from QG2083 by proteinase K digestion and isopropanol precipitation (Sunnucks and Hales 1996) and prepared paired-end (PE) 600-bp insert and mate-pair (MP) 4-kb insert Illumina Nextera libraries, from which 100-bp reads were sequenced to approximately 30x and 60x expected coverage (see below) on an Illumina HiSeq 2000 at the NYU Center for Genomics and Systems Biology GenCore facility (New York, USA). PE reads for QG2082 were generated similarly, for 20x expected coverage. PE reads for G_4BC_2 mapping populations were generated from 250bp insert libraries sequenced to mean 4x coverage per line with 150-bp reads on a HiSeq 2500. FASTX Toolkit (v. 0.0.13; options '-n -l 25 -M 9 -Q 33'; http://hannonlab.cshl.edu/fastx_toolkit/) was used to trim low quality bases from PE reads, and NextClip (v. 1.3; (Leggett *et al.* 2014)) was used to remove adaptors from MP reads and assess read orientation. Contaminating bacterial sequence was removed with Kraken (v. 1.0; (Davis *et al.* 2013)) using a database built from NCBI nt (downloaded 10/06/2014), and sequencing errors were corrected with Blue (v. 1.1.2; (Greenfield *et al.* 2014)). Raw data is available from the NCBI SRA under project PRJNA525787.

We assembled contigs and conservative scaffolds for QG2083 with the string graph assembler SGA (Simpson and Durbin 2012). Scaffolding required contigs to be linked by at least 5 unambiguously mapped mate-pair reads and full path consistency between contigs and all mapped mate-pair reads (option '--strict'). PacBio long reads (12x coverage, P4-C2 chemistry, Duke University Center for Genomic and Computational Biology) were used for a single round of gap closing and scaffolding with PBJelly (v. 13.10, minimum evidence of two reads; (English *et al.* 2012)), followed by scaffold breakage at any regions of mate-pair read discordance with REAPR (v. 1.0.17; (Hunt *et al.* 2013)). This yielded a fragmented assembly of 9152 scaffolds of length at least 500 bp spanning 95.9 Mb, with half the assembly in scaffolds of 56 kb or more.

QG2082 reads were aligned to the QG2083 reference scaffolds with bwa mem (v. 0.5.9; (Li and Durbin 2010)), processed with samtools (v. 1.2; http://www.htslib.org/) and Picard (v. 1.111; http://broadinstitute.github.io/picard/) utilities, and SNPs were called using the GATK HaplotypeCaller (v. 3.1-1; hard filtering on 'MQ < 40.0 || DP < 4 || FS > 60.0 || ReadPosRankSum < -8.0 || QD < 4.0 || DP > 80'; (McKenna *et al.* 2010)). Fixed diallelic SNPs were supplemented with calls from reference mapping of the pooled G_4BC_2 data and, to mitigate potential mapping bias against highly divergent regions, SNP calls from whole genome alignment of a *de novo* assembly of G_4BC_2 data pooled with QG2083 data at 1:1 expected heterozygosity, using the heterozygous aware assembler Platanus (Kajitani *et al.* 2013) and Mummer (Kurtz *et al.* 2004). This yielded a total of 1.63M diallelic SNPs across 6843 scaffolds spanning 89.2 Mb. These SNPs provide markers for genetic map construction.

We generated F_2 animals from reciprocal crosses between QG2083 and QG2082. We then performed 160 single-pair random-mating intercrosses among F_2s to generate G_3s, and

again among G_3s to generate G_4s. These intercrosses allow additional meioses to expand the genetic map. We backcrossed G_4 females to QG2082 males, generating progeny carrying one recombinant version of the genome and one intact QG2082 genome. To increase the quantity of DNA representing these genomes, we crossed females to QG2082 males again and allowed the resulting G_4BC_2 populations to grow for a single generation before DNA extraction and library preparation. Data for 87 G_4BC_2 populations with at least 0.5x expected coverage were mapped to the QG2083 scaffolds and genotyped by HMM using a modified MSG pipeline (assuming six crossovers per assembled genome, genotyping error rates of 0.001, and a scaling parameter on transition probabilities of $1x10^{-11}$, taking the dominant parental assignment as a single genotype for each scaffold) (Andolfatto *et al.* 2011). Markers were filtered in r/qtl based on missing data (<50%) and segregation distortion (χ^2 test $p > 1x10^{-10}$), followed by removal of redundant markers (Broman *et al.* 2003). Based on genotyping error rate estimates we further required a minimum of 70 SNP calls per scaffold and a minimum length of 1 kb, yielding 1337 non-redundant scaffold markers spanning 81.5 Mb. Six linkage groups were recovered over a wide range of minimum pairwise LOD scores, and within linkage group marker ordering was carried out by the minimum spanning tree method implemented in ASMap (Taylor and Butler 2017).

Linkage group scaffolds were aligned against multiple Platanus *de novo* assemblies using the sequencing data from inbred lines and subsampling of the approximately 350x pooled data from G_4BC_2 lines, with the UCSC chain/net pipeline (Kent *et al.* 2003). Genome evolution in the *Caenorhabditis* genus occurs largely through intrachromosomal rearrangement, and 84-94% of the net for each *C. becei* linkage group aligned to the homologous *C. elegans* chromosome. The assembly most concordant with genetic data was selected, based on the number and aligned

length of (1) scaffolds mapping to multiple linkage groups and (2) outliers in the ratio of genetic to physical distance. The selected assembly contained six cases where scaffolds mapped to multiple linkage groups – usually terminal regions of large scaffolds, with these terminal alignments spanning just over 1 Mb in total – and three scaffolds where mapping within linkage groups was inconsistent with the genetic data (median absolute deviation 99th percentile for the ratio of genetic to physical span), spanning 179 kb.

The draft assembly improved contiguity and incorporated several Mb of sequence not represented by the genetic map scaffolds, from 2483 redundant marker scaffolds to 256 final scaffolds spanning 84.9 Mb. We estimate a genome size of around 95 Mb based on mapped read-depth (Simpson and Durbin 2012), with 90 Mb of that unique sequence amenable to assembly and variant calling.

4.4.10 *C. nouraguensis genome assembly*

The *C. nouraguensis* reference genome was generated from JU2079, an inbred strain derived from 28 generations of JU1825 sibling matings (Marie-Anne Félix, personal communication). Genomic DNA and RNA were purified from mixed-stage individuals and sent to Mark Blaxter at the University of Edinburgh for Illumina sequencing. 125-bp paired-end reads were generated from two libraries, one with inserts of ~400 bp (~55 million read-pairs), the other with inserts of ~550 bp (~54 million read-pairs) (SRA project accession PRJEB10884). After generating a preliminary assembly, we used BlobTools (Laetsch and Blaxter 2017) for taxonomic classification. Some contigs matched *E. coli* (provided as food) and some matched an unsequenced bacterial species in the Firmicutes phylum. We extracted the Firmicutes contigs and used GSNAP (Wu and Nacu 2010) to remove matching read-pairs, and to remove any read-pairs

matching the *E. coli* REL606 genome (Genbank accession NC_012967), a strain closely related to OP50. We also removed reads failing Illumina's 'chastity filter' and used cutadapt to trim poor quality sequences (phred score < 10) and adapter sequences from the 3' ends of reads. We then performed error-correction using Musket (Liu *et al.* 2013) with a k-mer size of 28, and performed *de novo* genome assembly, scaffolding and gap closure using Platanus (Kajitani *et al.* 2013) with an initial k-mer size of 21. We then used REAPR (Hunt *et al.* 2013) to break any misassembled scaffolds. The resulting assembly is approximately 73 Mb in size. We used BUSCO (Simão *et al.* 2015) to measure its completeness using a set of conserved genes, and find that our assembly contains 81.5% of conserved single-copy orthologs as a single copy, 9.8% as duplicates, and 3.6% as fragments, very similar to what is seen in the finished *C. elegans* genome assembly (Simão *et al.* 2015). Using MUMMER (Kurtz *et al.* 2004), we ordered and oriented our *C. nouraguensis* scaffolds based on synteny to the *C. becei* genome assembly. Using this approach, some scaffolds remain unplaced and others may be misplaced, either because of true differences between the two species' genomes, or because MUMMER's alignments are noisy.

We also generated assemblies of the mitochondrial genomes of *C. becei* strain QG2083, and *C. nouraguensis* strain JU2079 using reads generated by Mark Blaxter's lab (EMBL ENA accession ERR1018617), filtered as above to remove *E. coli* reads, and down-sampled to 10 million read-pairs for *C. nouraguensis*. We then used the Assembly by Reduced Complexity (ARC) (Hunter *et al.* 2015) pipeline to assemble mitochondrial genomes, using the *C. elegans* mitochondrial genome sequence as a starting point.

4.4.11 *Whole-genome amplification and library preparation*

After testing the fertility of rare viable offspring from crosses of *C. nouraguensis* JU1825/NIC59 females to *C. becei* males (see above), we prepared Illumina whole-genome sequencing libraries as follows. We first transferred each worm individually to a blank NGM plate with no OP50 lawn for 30 minutes to reduce the amount of contaminating bacterial DNA. We then performed whole-genome amplification using the Qiagen REPLI-g Single Cell kit (Dey *et al.* 2012). We picked individual worms into 4 µl of PBS sc and froze them at -20° for at least an hour. We thawed the samples and added 3 µl of Buffer D2, mixed well, and incubated at 65° for 20 minutes rather than the recommended 10 minutes to aid in worm lysis. The rest of the protocol was performed according to the manufacturer's guidelines. We purified the resulting amplified DNA using the Zymo Research Genomic DNA Clean & Concentrator kit (gDDC-10). We submitted approximately 0.5-0.8 ng of each sample to the Fred Hutchinson Cancer Research Center Genomics Core Facility. A uniquely barcoded sequencing library for each sample was made using the Illumina Nextera library prep kit and sequenced on an Illumina Hiseq2500 to generate 50-bp paired-end reads. We sequenced eleven fertile and ten sterile F1 individuals. We also sequenced a set of control samples so that we could determine how coverage varies across the genome for pure *C. nouraguensis* or *C. becei* populations, as well as to allow us to determine fixed SNPs between the NIC59 and JU1825 *C. nouraguensis* strains. Specifically, we sequenced genomic DNA derived from large populations of NIC59, JU1825 and QG711 (sample names: NIC59_bulk, JU1825_bulk and QG711_bulk). In order to determine whether SNP frequencies and coverage metrics behave as expected in heterozygous or hybrid triploid genomes, we sequenced three additional controls: (a) a single heterozygous NIC59/JU1825 female (sample name: F1_NIC59_JU1825); (b) a sample where we placed one NIC59 female and one JU1825

female in the same tube (sample name: NIC59plusJU1825); and (c) a sample where we placed one NIC59 female, one JU1825 female and one QG711 female in the same tube (sample name: NIC59plusJU1825plusQG711).

4.4.12 *SNP calling, genotyping, and coverage calculations*

Using GSNAP (Wu and Nacu 2010), we mapped the reads from each sample to a combined reference genome containing the *C. nouraguensis* and *C. becei* nuclear and mitochondrial genome assemblies, the *E. coli* REL606 genome sequence and the Firmicutes contigs described above. We allowed GSNAP to report only a single map position for each read (--npaths=1). We further filtered read mappings, requiring mapq value of at least 20 to select uniquely mapping reads.

After preprocessing bam files by marking duplicates (using picard's MarkDuplicates, http://broadinstitute.github.io/picard/) and realigning indels (using GATK's RealignerTargetCreator and IndelRealigner tools (McKenna *et al.* 2010)), we called SNPs in *C. nouraguensis* scaffolds using samtools mpileup (Li *et al.* 2009) (ignoring indels, disabling the per-base alignment quality option and increasing the depth down-sampling parameter to 6660). We counted reads matching each allele using GATK's VariantAnnotator tool (McKenna *et al.* 2010), disabling the down-sampling option and using the "ALLOW_N_CIGAR_READS" option. We noticed that the distribution of allele frequencies in our 'heterozygous' control samples was skewed towards the reference allele, likely because those reads are easier to map to the reference assembly. We therefore used this first round of SNP calls to remap all reads to the combined reference genome using GSNAPs "SNP-tolerant" mode that allows reads to map to either haplotype. SNPs were called a second time, after one additional level of read mapping

filtering, where we used the R/Bioconductor Rsamtools package (Huber *et al.* 2015) to select only reads where the full length of the read could be aligned to the reference genome.

We then used the R/Bioconductor VariantAnnotation package (Huber *et al.* 2015) to select high-quality fixed differences between the NIC59 and JU1825 strains as follows: we required a SNP quality score of at least 100; that opposite alleles be fixed (non-reference frequencies of <5% and >95%) in the two strains; that read depth in some *C. nouraguensis* control samples be within typical range for that sample (5-50 for JU1825_bulk and NIC59_bulk, and 100-450 for the combined JU2079 samples); and that read depth be 0 in our *C. becei* control samples (QG711_bulk and the two QG2083 libraries). 337,493 SNPs passed these filters (an average of one SNP every 217 bp). To estimate mean allele frequencies in 50-kb windows across the genome, we counted the total number of reads matching NIC59 and JU1825 alleles across the window and divided the NIC59 count by the total count. We then used a circular binary segmentation algorithm, implemented in the Bioconductor package DNAcopy (Olshen *et al.* 2004), to estimate the locations of breaks between different haplotypes as well as the average allele frequency in each segment.

Using the same filtered bam files, we determined coverage at each base position using the samtools depth tool (Li *et al.* 2009). Examining coverage and aligned bases in our control samples (JU1825_bulk, NIC59_bulk and QG711_bulk) reveals very high within-species sequence diversity, as seen in some other nematodes (Cutter *et al.* 2006, 2013; Dey *et al.* 2013). This high diversity means that in some genomic regions, especially on the chromosome arms, short reads are unalignable to the reference genomes, resulting in many regions of no coverage. Other genomic regions show more aligned base-pairs, but still harbor many SNPs. Therefore, before calculating mean coverage metrics, we filtered the base positions under consideration to

include only those that had coverage of at least 8 reads in the relevant control samples (JU1825_bulk and NIC59_bulk for the *C. nouraguensis* assembly and QG711_bulk for the *C. becei* assembly), and then determined mean values in 50-kb regions of each scaffold. This filtering mitigated diversity-related coverage variation a little, but even in control samples coverage is still lower on chromosome arms than centers. To generate per-chromosome copy number estimates (Fig. 4A), we manually defined 1-2 regions in the center of each chromosome that are well-covered in each of the control samples and took the mean of 50-kb window coverages in each region.

4.4.13 *Fixing and staining embryos*

60-80 gravid females were picked into 30 µl of 1x Egg Buffer + 0.1% Tween-20 on a glass coverslip and dissected in half with needles to release embryos. After the females were dissected, we waited for five minutes to allow the mostly 1-cell stage embryos to develop. A Histobond slide was then gently laid on top of the coverslip. Excess liquid was removed with Whatman paper and the samples were frozen by placing the slide on a metal block on dry ice for 10 minutes. Coverslips were then quickly flipped off with a razor blade and the slides were immediately placed in a Coplin jar with -20° methanol for 10 min. The slides were then moved to a Coplin jar with -20° acetone for 10 min. The slides were washed three times for 10 mins each in Coplin jars with 1x PBS + 0.1% Tween-20, then blocked for 30 mins in a Coplin jar with 1x PBS + 0.1% Tween-20 + 0.5% BSA + 0.02% sodium azide (blocking buffer). Each slide was then incubated overnight at 4° with a 1:100 dilution of mouse monoclonal anti-α-tubulin primary antibody (clone DM1A, Sigma) in blocking buffer. The next day, the slides were washed three times for 10 mins each in Coplin jars with 1x PBS + 0.1% Tween-20. Each slide was then

incubated in the dark for 2 hours at room temperature with a 1:1000 dilution of Alexa-Fluor 488 goat anti-mouse secondary antibody (Jackson ImmunoResearch) diluted in 1x PBS + 0.1% Tween-20. The slides were moved to a Coplin jar with 1x PBS + 0.1% Tween-20 for 10 minutes, then a Coplin jar with 1x PBS + 0.1% Tween-20 + 5 µl of 5 mg/ml DAPI for 15 minutes, then a Coplin jar with 1x PBS + 0.1% Tween-20 for 20 minutes. Samples were mounted with 12 µl VectaShield (H-1000, Vector Laboratories) and sealed with nail polish. Samples were imaged using a Nikon Eclipse 80i compound microscope with a 100x oil objective (1.45 NA). Images were processed using Image J.

We only analyzed embryos at early stages of embryogenesis (1-4 cell stage). In normal embryos at this stage, the DNA in the two polar bodies appears as two distinct clumps at the periphery of the embryo that are easily distinguishable from other embryonic nuclei. Polar bodies in hybrid embryos were considered to have abnormal morphologies if they were unlike those observed in the JU1825 control. This encompasses a range of abnormalities such as not being round, being larger, or having easily distinguishable chromatids.

4.4.14 *Male UV irradiation*

Virgin *C. nouraguensis* L4 males and L4 females were collected and separately placed onto two new NGM plates seeded with OP50 and rimmed with palmitic acid. We added a few females to the male plate to coax the males to stay on the surface of the plate. The next day, now young adult males were picked onto a blank NGM plate rimmed with palmitic acid. This plate was placed in a CL-1000 UV Crosslinker (Ultra-Violet Products) without a lid and the males were exposed to 70,000 µJ/cm^2 254-nm UV radiation. Immediately afterward, the males were picked onto new NGM plates seeded with OP50 and rimmed with palmitic acid. Each plate had

thirty irradiated males and thirty young adult virgin females. The next day, rare viable larvae were picked from the cross plates onto a new NGM plate seeded with OP50 and rimmed with palmitic acid. When these progeny reached adulthood, they were individually frozen, lysed and PCR genotyped.

We tested the following UV exposures: 15,000, 40,000, 50,000, 60,000, 70,000, 80,000, 150,000 and 250,000 μJ/cm^2. Males treated with lower exposures (15,000 and 40,000 μJ/cm^2) were able to produce many viable progeny, indicating that their sperm DNA was not damaged enough to induce paternal genome loss. Males treated with the highest exposures (150,000 and 250,000 μJ/cm^2) mostly stopped moving and never recovered. We found that males treated with 70,000 μJ/cm^2 were healthy enough to mate and produce many dead fertilized embryos, but relatively few viable progeny.

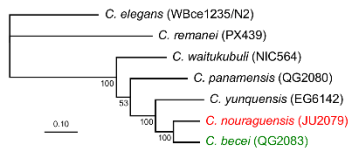
4.5 ACKNOWLEDGEMENTS

We thank Janet Young for assembling the *C. nouraguensis* (JU2079). We thank Marie-Anne Félix and Christian Braendle for providing all the strains of *C. nouraguensis* used in this study; Max Bernstein, Vicky Cattani, Taniya Kaur, Jasmine Nicodemus and Annalise Paaby for help performing *C. becei* mapping crosses and preparing sequencing libraries; Mark Blaxter, Lewis Stevens and members of the *Caenorhabditis* Genomes Project for sequencing *C. nouraguensis* (JU2079) as well as providing a public BLAST server of newly sequenced *Caenorhabditis* species (http://blast.caenorhabditis.org/); Maulik Patel for help in preparing DNA for *C. nouraguensis* genome sequencing (JU1825 and NIC59); Jihong Bai for sharing his UV-irradiator; Lamia Wahba for advice on single-worm sequencing; Needhi Bhalla and Mara Schvarzstein for discussions of *C. elegans* female meiosis; Jim Priess for discussions of *C.*

elegans development; and Hannah Seidel for advice on time-lapse imaging of *Caenorhabditis* embryos. We also thank Barbara Wakimoto, Diane Shakes, Lisa Kursel, Irini Topalidou, Ofer Rog and members of his lab, and members of the Ailion lab for comments on the manuscript. We thank the Fred Hutchinson Genomics core for Illumina library preparation and sequencing. P.L. was supported in part by a National Institutes of Health Institutional Training Grant (Public Health Service, National Research Service Award, T32GM007270 from National Institute of General Medical Sciences). This work was supported by the NYU CGSB Genomics Core facility and NYU IT High Performance Computing resources, NIH grant R01 GM121828 to M.V.R, NIH grant R01 GM074108 to H.S.M, and an NSF CAREER Award (MCB-1552101) to M.A.

4.6 FIGURES

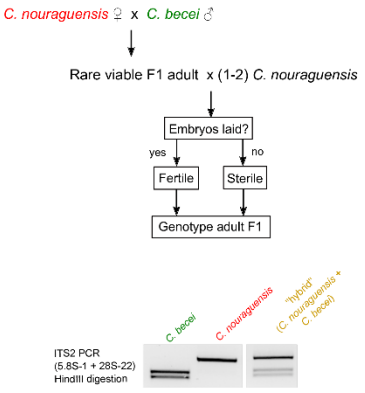
A



B

cross	viable F1 adults	dead F1 embryos	frequency of viable F1 adults (%)
C. nouraguensis (JU1825) ♀ x *C. becei* (QG711) ♂	5	6239	0.08
C. nouraguensis (NIC59) ♀ x *C. becei* (QG711) ♂	2	5913	0.03
C. nouraguensis (NIC54) ♀ x *C. becei* (QG711) ♂	14	4227	0.33
C. becei (QG711) ♀ x *C. nouraguensis* (JU1825) ♂	0	4530	0
C. becei (QG711) ♀ x *C. nouraguensis* (NIC59) ♂	0	3091	0
C. becei (QG711) ♀ x *C. nouraguensis* (NIC54) ♂	0	4308	0

C



D

C. nouraguensis (JU1825) ♀ x *C. becei* (QG711) ♂

Genotype of F1 adult at W02B12.9 locus (Chr. II)

		Hybrid (*C. nouraguensis*/*C. becei*)		Maternal (*C. nouraguensis*)		
		female	male	female	male	total
Fertility of F1 adult	Fertile	0	0	7	1	8
	Sterile	13	4	1	0	18
	total	13	4	8	1	26

C. nouraguensis (NIC54) ♀ x *C. becei* (QG711) ♂

Genotype of F1 adult at ITS2 locus (Chr. I)

		Hybrid (*C. nouraguensis*/*C. becei*)		Maternal (*C. nouraguensis*)		
		female	male	female	male	total
Fertility of F1 adult	Fertile	0	0	10	0	10
	Sterile	21	5	1	0	27
	total	21	5	11	0	37

Figure 4.1. Crossing *C. nouraguensis* females to *C. becei* males results in sterile F1 with hybrid genotypes and fertile F1 with only maternal genotypes.

(A) A maximum likelihood phylogeny of several *Caenorhabditis* species closely related to *C. nouraguensis* and *C. becei*, with strain names in parentheses. *C. elegans and C. remanei* were used as outgroups. The scale bar represents 0.10 substitutions per site. Bootstrap support values are indicated to the left of each node (percent of 100 bootstrap replicates). See Stevens et al. (2018) for a more complete *Caenorhabditis* phylogeny (Stevens *et al.* 2018). **(B)** Frequency of viable F1 adults in crosses between *C. nouraguensis* and *C. becei*. NIC54 females yield a significantly higher proportion of viable F1 adults when crossed to *C. becei* (QG711) males than either JU1825 or NIC59 females (Chi-square with Yates correction, P=0.0065 NIC54-JU1825, P=0.0005 NIC54-NIC59). There was no significant difference in the frequency of rare viable F1 adults produced by JU1825 and NIC59 females (Chi-square with Yates correction, P=0.49). The

crosses between *C. becei* females and *C. nouraguensis* males serve as controls for accidental contamination of plates with embryos or larvae from either parental species: no viable adults were found among more than 12,000 F1 screened (Figure 4.9). **(C)** Flowchart showing how fertility of rare viable F1 was tested (also see Materials and Methods). The gel shows how a PCR-RFLP assay distinguishes between *C. nouraguensis* and *C. becei* alleles at the ITS2 locus. **(D)** Tables showing the relationship between the fertility, genotype and sex of rare viable F1 derived from crossing either *C. nouraguensis* (JU1825) females to *C. becei* (QG711) males (top table, genotyped at the W02B12.9 locus) or *C. nouraguensis* (NIC54) females to *C. becei* (QG711) males (bottom table, genotyped at the ITS2 locus). Both sterile and fertile F1 exhibited a more strongly female-biased sex ratio than that seen in intraspecies crosses (Figure 4.8). See also Figures 4.8-4.10 and Videos S1-S4.

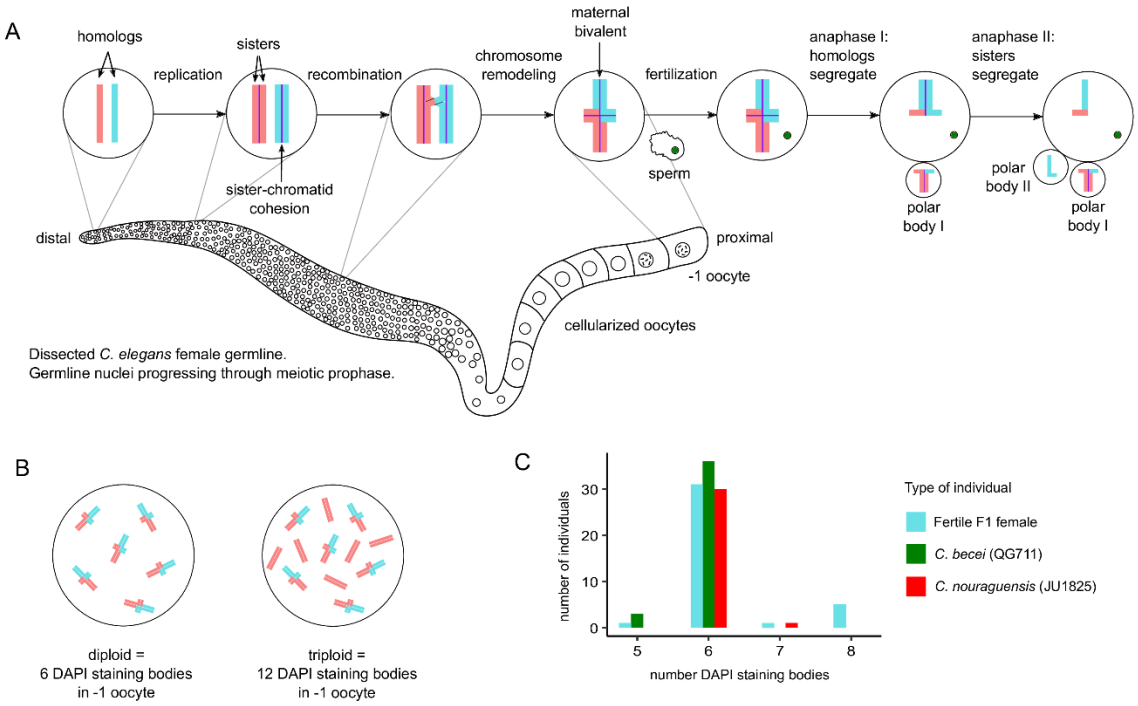
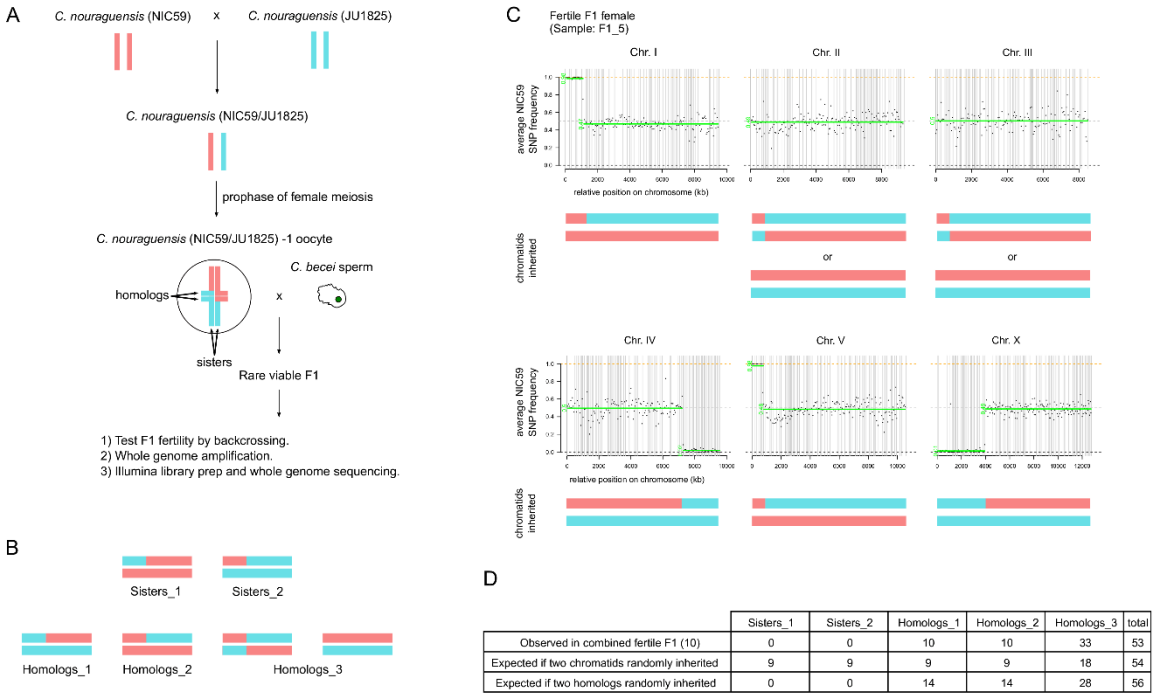


Figure 4.2. Fertile interspecific F1 females are diploid.

(A) Schematic of one of the two germlines of a *Caenorhabditis* female. The germline is a syncytial tube with nuclei (depicted as small circles) hugging its circumference. Germline nuclei are generated from mitotically dividing stem cells at the distal tip and migrate proximally while undergoing meiotic prophase. Homologs of each of the six chromosomes replicate their DNA, with the resulting sister chromatids held together by sister chromatid cohesion (purple lines). Homologous chromosomes undergo one crossover biased towards a chromosome end and are remodeled into a cruciform structure. Eventually, cell membranes form around the nuclei, resulting in mature oocytes. In mature oocytes, homologous chromosomes and their sister chromatids form bivalents that are held together by a combination of sister chromatid cohesion and recombination. The -1 oocyte is fertilized by sperm, triggering the reductional first meiotic division in which sister chromatid cohesion is lost between homologs, allowing them to segregate during anaphase I. One set of homologs is segregated into the first polar body. During

the second meiotic division, sister chromatid cohesion is lost between sister chromatids, which then segregate during anaphase II. One chromatid is segregated into the second polar body and the other is inherited by the oocyte. **(B)** Because only one crossover occurs per homologous set of chromosomes, diploid individuals have six bivalents (6 DAPI-staining bodies) while triploids have six bivalents and six univalents (12 DAPI-staining bodies). **(C)** Most *C. nouraguensis* (JU1825), *C. becei* (QG711) and fertile F1 females have 6 DAPI-staining bodies, consistent with being diploid for six chromosomes. However, some fertile F1 females have a slightly higher number of DAPI-staining bodies, suggesting they are mostly diploid, but inherit some extra pieces of DNA. See also Figure 4.11.

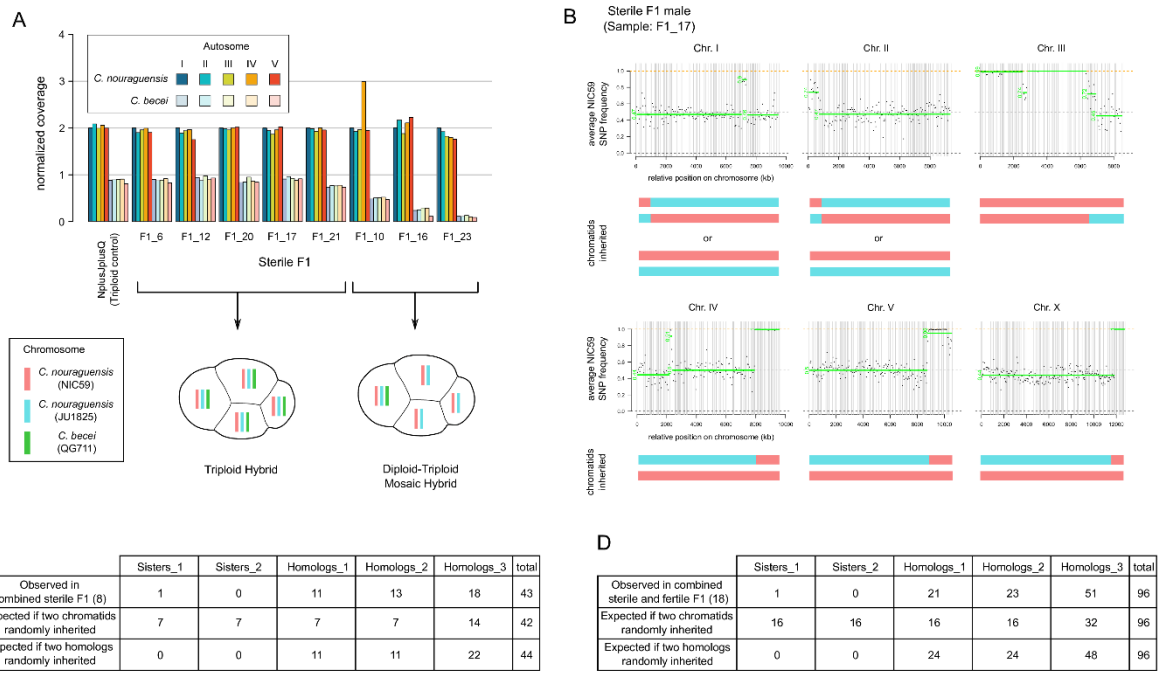


	Sisters_1	Sisters_2	Homologs_1	Homologs_2	Homologs_3	total
Observed in combined fertile F1 (10)	0	0	10	10	33	53
Expected if two chromatids randomly inherited	9	9	9	9	18	54
Expected if two homologs randomly inherited	0	0	14	14	28	56

Figure 4.3. Fertile interspecific F1 inherit two randomly selected homologous chromatids from each maternal bivalent.

(A) Schematic of how we determined which two chromatids are inherited from each maternal *C. nouraguensis* bivalent. Two genetically distinct strains of *C. nouraguensis*, NIC59 and JU1825, were crossed to make heterozygous NIC59/JU1825 females, which were then crossed to *C. becei* (QG711) males. Viable F1 progeny were collected, assayed for fertility by backcrossing, and prepared for whole-genome amplification and sequencing. **(B)** The six possible ways to inherit two chromatids from a bivalent. Four have distinct chromosomal genotypes (Sisters_1, Sisters_2, Homologs_1 and Homologs_2), and two have genotypes that cannot be distinguished in our sequencing data (Homologs_3, which are heterozygous everywhere). **(C)** An example of whole-genome genotyping data from a single fertile F1 female (F1_5). Each plot represents one of the six chromosomes. Each point represents the average NIC59 SNP frequency in 50-kb windows ordered along the chromosome; haplotype change points and average allele frequency for each

segment are shown by the green horizontal lines. The reference assembly is fragmented into scaffolds that we ordered and oriented on chromosomes using synteny; gray vertical lines represent breaks between scaffolds. The combination of two chromatids that best matches the genotyping data is shown underneath each plot. Genotyping and coverage plots for all F1 are shown in Figures 4.12 and 4.13. One fertile F1 gave ambiguous genotypes and was excluded from further analysis (F1_25); this individual had a particularly high percentage of contaminating bacterial reads and low coverage of the *C. nouraguensis* genome (Table 4.1). **(D)** Frequency of the five distinguishable chromosome genotypes when combining all fertile F1 genotyping data. The table also shows the expected frequency of the five genotypes if any two chromatids were randomly inherited and if any two homologous chromatids were inherited. The frequencies observed in the fertile F1 are not different from those expected from random inheritance of two homologous chromatids (Fisher's exact test, P=0.41). Hemizygous X chromosomes in males and triploid autosomes were excluded from the analysis. See also Figures 4.12 and 4.13 and Table 4.1.



C

	Sisters_1	Sisters_2	Homologs_1	Homologs_2	Homologs_3	total
Observed in combined sterile F1 (8)	1	0	11	13	18	43
Expected if two chromatids randomly inherited	7	7	7	7	14	42
Expected if two homologs randomly inherited	0	0	11	11	22	44

D

	Sisters_1	Sisters_2	Homologs_1	Homologs_2	Homologs_3	total
Observed in combined sterile and fertile F1 (18)	1	0	21	23	51	96
Expected if two chromatids randomly inherited	16	16	16	16	32	96
Expected if two homologs randomly inherited	0	0	24	24	48	96

Figure 4.4. Sterile interspecific F1 inherit a diploid *C. nouraguensis* genome and a haploid *C. becei* genome.

(A) The normalized read coverage for all autosomes in eight sterile F1 with a clear genotype. Coverage is normalized to *C. nouraguensis* chromosome I coverage in each individual, which is set to two. Schematics below the graph show sterile F1 that are fully triploid hybrids (all cells of an F1 embryo have a diploid *C. nouraguensis* genome and a haploid *C. becei* genome) or diploid-triploid mosaic hybrids (all cells of an F1 embryo have a diploid *C. nouraguensis* genome, but only a subset of cells inherit the haploid *C. becei* genome). **(B)** An example of whole genome *C. nouraguensis* genotyping data from a single sterile F1 male (F1_17). Each plot represents one of the six chromosomes. Each point represents the average NIC59 SNP frequency (after removing *C. becei* reads) in 50-kb windows along the physical length of the chromosome; haplotype change points and average allele frequency for each segment are shown by the green horizontal lines. The gray vertical lines represent breaks between scaffolds. The combination of two *C. nouraguensis* chromatids that best matches the genotyping data is shown underneath each

plot. Genotyping and plots for all sterile F1 are shown in Figures 4.12 and 4.13. We excluded two sterile F1 from further analysis because they gave ambiguous genotypes (F1_18 and F1_26). **(C)** Frequency of the five distinguishable *C. nouraguensis* chromosome genotypes when combining all sterile F1 genotyping data. The table also shows the expected frequency of the five genotypes if any two chromatids were randomly inherited and if any two homologous chromatids were inherited. The frequencies observed in the sterile F1 are not different from those expected from random inheritance of two homologous chromatids (Fisher's exact test, P=0.78). **(D)** Frequency of *C. nouraguensis* chromosome genotypes when combining all fertile and sterile F1 genotyping data. The frequencies observed are not different from those expected from random inheritance of two homologous chromatids (Fisher's exact test, P=0.85). See also Figures 4.12 and 4.13 and Table 4.1.

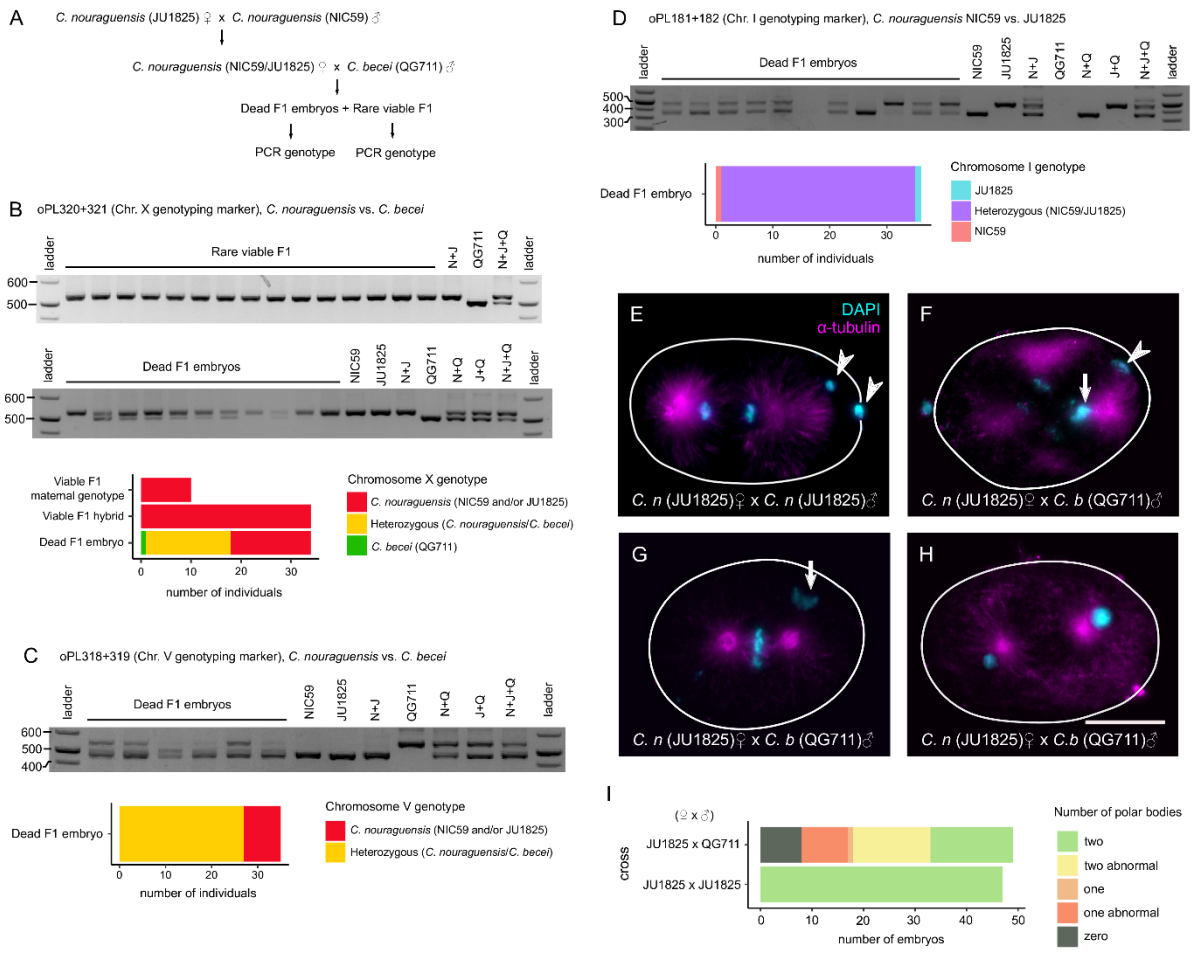


Figure 4.5. Dead interspecific F1 embryos inherit the *C. becei* X-chromosome and at least two maternal homologous chromatids.

(A) Schematic illustrating how dead F1 embryos and viable F1 progeny were generated for PCR genotyping experiments. **(B)** An example of a DNA gel showing the chromosome X genotypes (oPL320+321) of rare viable F1 adults (top gel) and dead F1 embryos (bottom gel). The graph shows that no viable F1 animals inherited the *C. becei* X-chromosome, but half of the dead embryos inherited it. **(C)** An example of a DNA gel showing the chromosome V genotypes (oPL318+319) of dead F1 embryos. The graph shows that most dead F1 embryos inherited chromosome V from both parents but that some inherited only a maternal (*C. nouraguensis*) copy. **(D)** An example of a DNA gel showing the chromosome I genotypes (oPL181+182) of

dead F1 embryos. Importantly, the primers used in this PCR reaction do not produce products when using a control *C. becei* QG711 embryo lysate as template, so any signal should be from *C. nouraguensis* templates. The graph shows that almost all dead F1 embryos have a heterozygous NIC59/JU1825 genotype. **(E)** A one-cell stage embryo derived from crossing *C. nouraguensis* (JU1825) females to *C. nouraguensis* (JU1825) males, fixed and stained for DNA (cyan) and alpha-tubulin (magenta). Tubulin staining shows the mitotic spindle, which helps determine how many cells there are in the embryo and at what point those cells are in the cell cycle. The white lines outline cell membranes. The two polar bodies that remain associated with the embryo are indicated by the arrowheads. **(F)** A hybrid embryo derived from crossing *C. nouraguensis* (JU1825) females to *C. becei* (QG711) males. One normal polar body is indicated by the arrowhead. An abnormally structured polar body (not round) is indicated by the arrow. **(G)** A hybrid embryo with only a single abnormal polar body (large and not round) indicated by the arrow. **(H)** A hybrid embryo with zero polar bodies. Scale bar: 50 μm. **(I)** Hybrid embryos can have fewer than two and abnormal polar bodies. By contrast, embryos derived from intraspecies JU1825 crosses always have two polar bodies. “Two abnormal” refers to embryos that have two polar bodies, but one or both have an abnormal structure. “One abnormal” refers to embryos that have a single polar body with an abnormal structure.

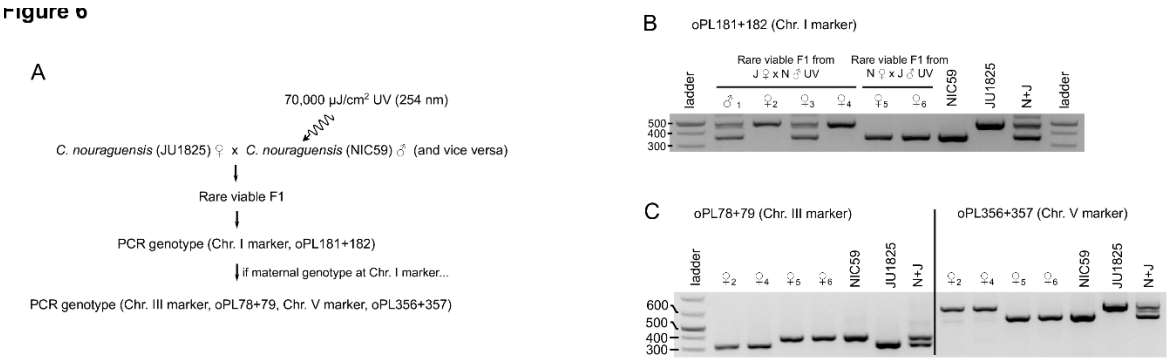


Figure 4.6. Diploid maternal inheritance can occur independently of interspecies hybridization.

(A) Flowchart illustrating how the UV irradiation experiments were conducted. **(B)** A DNA gel showing the chromosome I genotypes (oPL181+182) of several rare viable F1 derived from crossing either *C. nouraguensis* JU1825 females to UV-irradiated *C. nouraguensis* NIC59 males (J♀ x N♂ UV) or NIC59 females to UV-irradiated JU1825 males (N♀ x J♂ UV). The sex of the rare viable F1 is depicted above each lane, with an identifying number as subscript. **(C)** A DNA gel showing the chromosome III and V genotypes (oPL78+79 and oPL356+357) of the rare F1 from Figure 4.6B that had a maternal genotype for chromosome I. All also had a maternal genotype at these two markers.

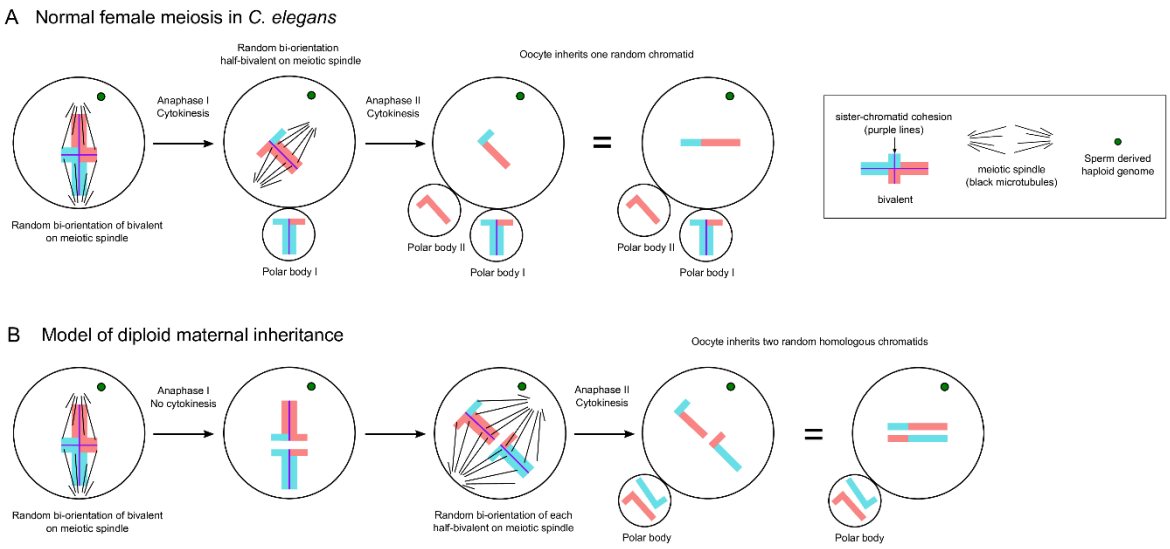


Figure 4.7. Model for diploid maternal inheritance in *C. nouraguensis* oocytes.

(A) Schematic of canonical meiotic divisions in *C. elegans*. Upon fertilization, each maternal bivalent (only one shown here) randomly bi-orients its homologs on the meiotic spindle, cohesion is lost between homologous chromosomes and homologs segregate. One set of homologs segregates into the first polar body while the other is retained in the oocyte (Anaphase I and Cytokinesis). Then the half-bivalent in the oocyte randomly bi-orients on the meiotic spindle, sister chromatid cohesion is lost and sister chromatids segregate. One sister chromatid is segregated into the second polar body while the other is retained in the oocyte (Anaphase II and Cytokinesis). Thus, the oocyte inherits only one random chromatid from each bivalent. **(B)** One model for how female meiosis could be modified to inherit two random homologous chromatids from a bivalent. Upon fertilization, the bivalent randomly bi-orients its homologs on the meiotic spindle and cohesion is lost between homologous chromosomes as is normal. Homologs segregate but cytokinesis fails (Anaphase I) and both half-bivalents remain in the oocyte. Each half-bivalent then bi-orients on the meiotic spindle, sister chromatid cohesion is lost, and sister chromatids segregate. One chromatid from each half-bivalent segregates into the second polar

body while the other is retained in the oocyte (Anaphase II and Cytokinesis). Thus, the oocyte inherits two random homologous chromatids from a bivalent.

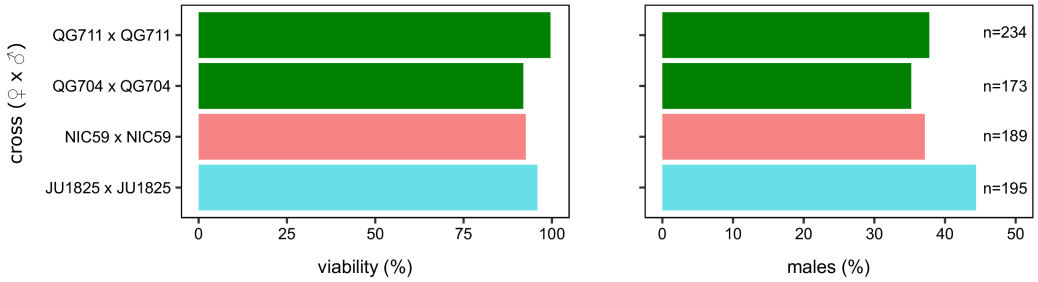


Figure 4.8. Control crosses to measure intra-strain viability. Related to Figure 4.1.

Wild isolates of *C. becei* (QG704 and QG711) and *C. nouraguensis* (NIC59 and JU1825) have high levels of intra-strain viability. All strains have a sex ratio skewed towards females, some of which are statistically significant (Fisher's exact test with Bonferroni correction, JU1825 p=1.0, NIC59 p=0.06, QG711 p=0.03, QG704 p=0.03).

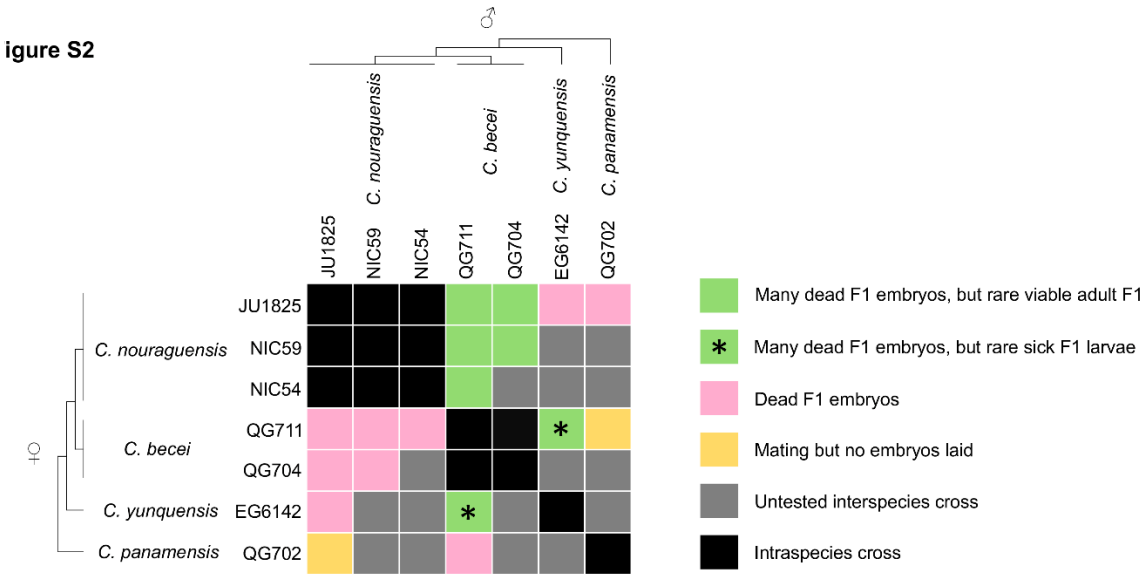


Figure 4.9. Summary of interspecies crosses. Related to Figure 4.1.

Rows show the females of each cross while males are shown in columns. The wild isolate strains used for each species are indicated. Black boxes are intraspecies crosses. Grey boxes are untested interspecies hybridizations. Rare viable F1 adults are present only when crossing *C. nouraguensis* females to *C. becei* males. Rare viable but sick F1 larvae are present in both directions of *C. becei* x *C. yunquensis* crosses. Worms mate but do not produce F1 embryos in *C. panamensis* female x *C. nouraguensis* male and *C. becei* female x *C. panamensis* male crosses. At least 12,000 dead F1 were screened for each cross that gave embryos.

A

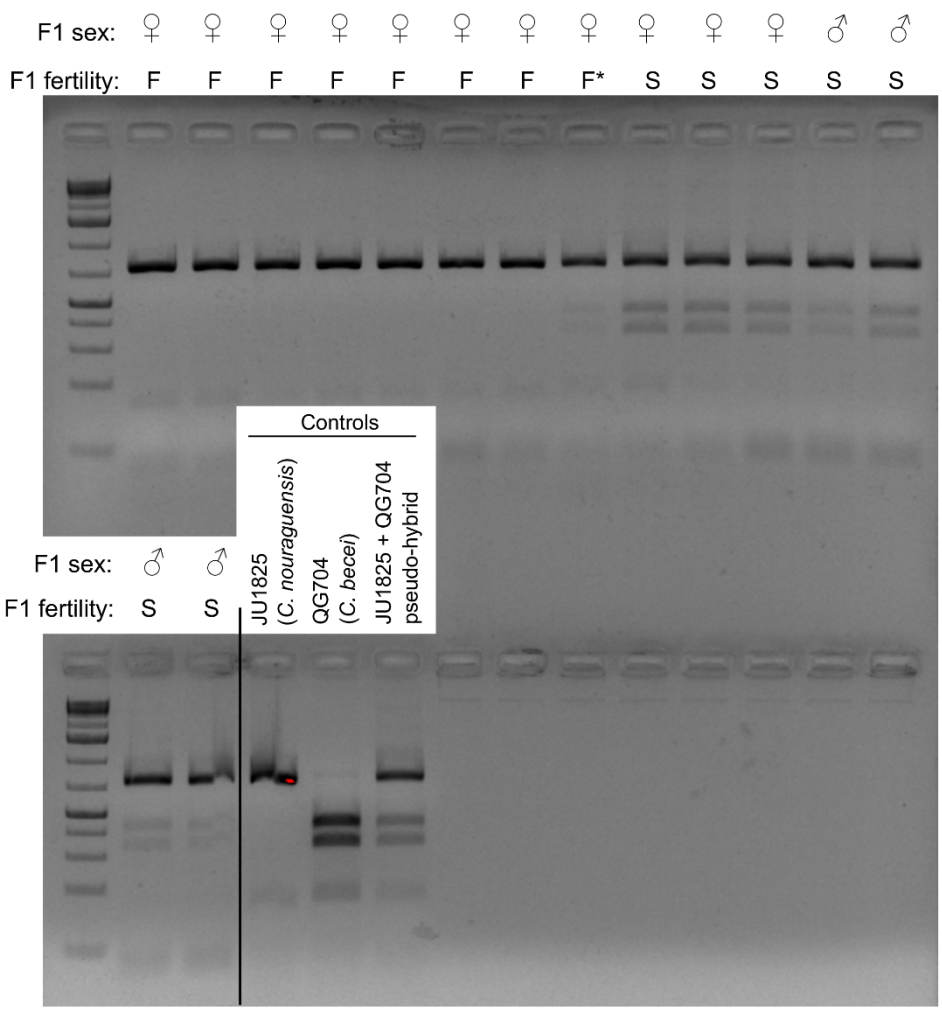
C. nouraguensis (JU1825) ♀ x *C. becei* (QG704) ♂

↓

Rare viable F1

1) Backcross F1 to JU1825 to test fertility (F= fertile, F*=fertile but dead F2 progeny, S=sterile).
2) PCR genotype F1 at ITS2 locus (Primers= 5.8S-1 + 28S-22, HindIII-HF digest).

B



C

Fertility of F1 adult	Hybrid (*C. nouraguensis*/*C. becei*)		Maternal (*C. nouraguensis*)		
	female	male	female	male	total
Fertile	1	0	7	0	8
Sterile	3	4	0	0	7
total	4	4	7	0	15

Figure 4.10. Fertile progeny with a maternal genotype and sterile progeny with a hybrid genotype are produced when C. nouraguensis females are crossed to males of a different C. becei strain. Related to Figure 4.1.

(A) We tested the fertility of rare viable F1 derived from *C. nouraguensis* JU1825 x *C. becei* QG704 crosses by backcrossing to JU1825 individuals of the opposite sex. F1 were then genotyped at the ITS2 locus using a PCR-RFLP assay. **(B)** A gel showing the sex, fertility and genotype at the ITS2 locus of viable adult F1. All fertile F1 have a maternal genotype, with one exception: one hybrid F1 female laid inviable F2 embryos (F*, still considered fertile). All sterile F1 had a hybrid genotype. **(C)** A table summarizing the genotyping data in (B).

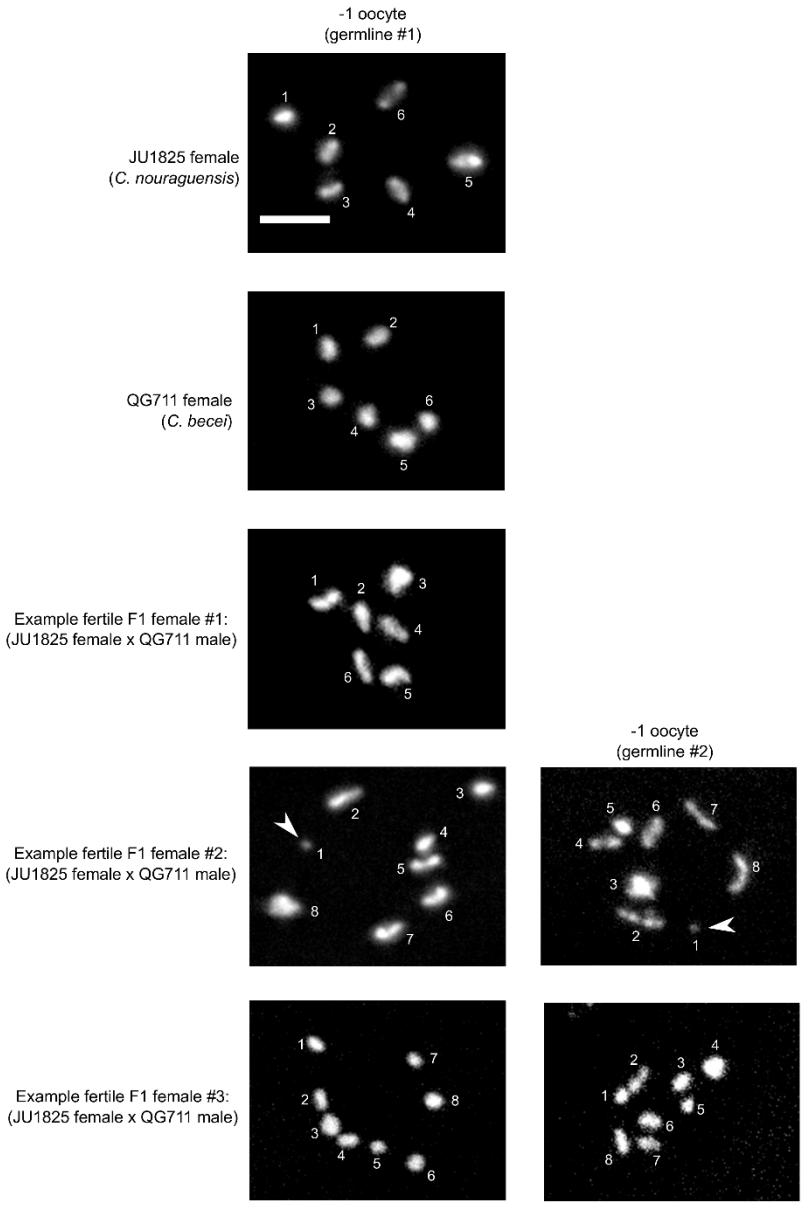
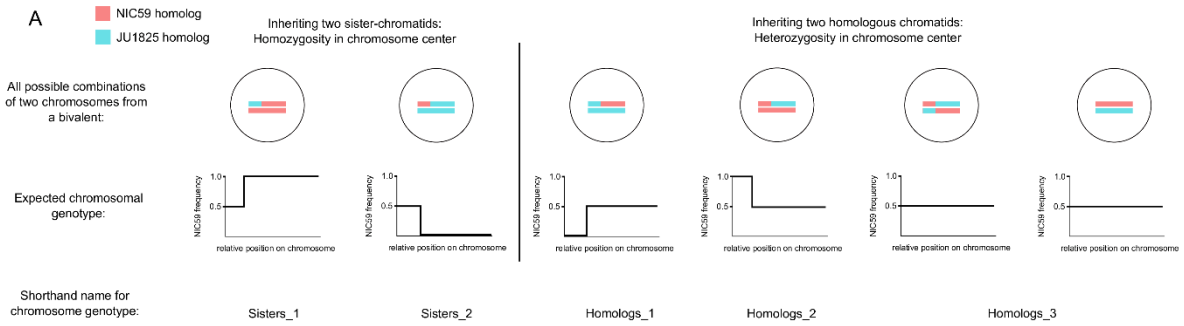


Figure 4.11. Fertile F1 females are diploid. Related to Figure 4.2.

The -1 oocytes from *C. nouraguensis* JU1825 and *C. becei* females primarily have six DAPI-staining bodies (two examples shown here). Most fertile F1 females derived from JU1825 female x QG711 male crosses have six DAPI-staining bodies (Example fertile F1 female #1). A minority have eight DAPI-staining bodies. In three of these cases, there appear to be seven relatively normal sized DAPI bodies plus a very small one (Example fertile F1 female #2, small

DAPI body highlighted by white arrowhead). In the other two cases, all eight DAPI bodies appear roughly equal in size (Example fertile F1 female #3). This higher number of DAPI-staining bodies is not the chance observation of a low frequency meiotic defect in a nucleus that happens to be in the -1 oocyte position (for example, homologs fail to recombine and increase the number of univalents) because we observed the same number of DAPI-staining bodies in both germlines of the same fertile F1 female. We hypothesize that these extra DAPI-staining bodies represent extra DNA (either maternal or paternal) in addition to the two chromatids inherited from each maternal bivalent. Scale bar: 5 μm.



B

C. nouraguensis chromosome genotypes

F1 sample name	fertility	sex	inferred ploidy	Chr. I	Chr. II	Chr. III	Chr. IV	Chr. V	Chr. X
F1_1	fertile	female	diploid	Homologs_3	Homologs_3	Homologs_3	Homologs_2	Homologs_2	Homologs_2
F1_5	fertile	female	diploid	Homologs_2	Homologs_3	Homologs_3	Homologs_1	Homologs_2	Homologs_1
F1_29	fertile	female	diploid	Homologs_3	Homologs_ambiguous	Homologs_3	Homologs_1	Homologs_3	Homologs_3
F1_41	fertile	female	diploid	Homologs_2	Homologs_3	Homologs_3	Homologs_3	Homologs_3	Homologs_3
F1_8	fertile	female	likely diploid, backcross contamination	Homologs_3	Homologs_3	Homologs_3	Homologs_3	Homologs_3	Homologs_2
F1_11	fertile	female	likely diploid, backcross contamination	Homologs_3	Homologs_3	Homologs_3	Homologs_1	Homologs_1	Homologs_3
F1_39	fertile	female	likely diploid, backcross contamination	Homologs_3	Homologs_1	Homologs_3	Homologs_1	Homologs_3	Homologs_1
F1_25	fertile	female	ambiguous, backcross contamination	ambiguous	ambiguous	ambiguous	ambiguous	ambiguous	ambiguous
F1_46	fertile	male	diploid	Homologs_ambiguous	Homologs_3	Homologs_2	Homologs_1	Homologs_3	Hemizygous X
F1_4	fertile	male	diploid	Homologs_2	Homologs_3	Homologs_3	Triploid	Homologs_3	Hemizygous X
F1_48	fertile	male	diploid	Homologs_2	Homologs_3	Triploid	Homologs_1	Homologs_3	Hemizygous X
F1_26	sterile	female	ambiguous, backcross contamination	ambiguous	ambiguous	ambiguous	ambiguous	ambiguous	ambiguous
F1_18	sterile	female	ambiguous, backcross contamination	ambiguous	ambiguous	ambiguous	ambiguous	ambiguous	ambiguous
F1_6	sterile	female	triploid hybrid	Homologs_2	Homologs_3	Homologs_3	Homologs_3	Homologs_1	Homologs_1
F1_12	sterile	female	triploid hybrid	Homologs_1	Homologs_ambiguous	Homologs_3	Homologs_3	Homologs_3	Homologs_3
F1_20	sterile	female	triploid hybrid	Homologs_2	Homologs_2	Homologs_3	Homologs_3	Homologs_3	Homologs_1
F1_16	sterile	male	diploid-triploid hybrid	Homologs_1	Homologs_1	Homologs_1	Homologs_3	Homologs_3	Hemizygous X
F1_23	sterile	male	diploid-triploid hybrid	Homologs_3	Homologs_2	Homologs_1	Homologs_1	Sisters_3	Hemizygous X
F1_10	sterile	male	diploid-triploid hybrid	Homologs_2	Homologs_2	Homologs_3	Triploid	Homologs_3	Homologs_2
F1_17	sterile	male	triploid hybrid	Homologs_3	Homologs_3	Sisters_1	Homologs_2	Homologs_2	Homologs_2
F1_21	sterile	male	triploid hybrid	Homologs_3	Homologs_1	Homologs_1	Homologs_2	Homologs_2	Homologs_2

C

Fertile F1

Chromosome genotype	Chr. I	Chr. II	Chr. III	Chr. IV	Chr. V	Chr. X	total
Homologs_1	0	1	0	6	1	2	10
Homologs_2	4	0	1	1	2	2	10
Homologs_3	5	8	8	2	7	3	33
Homologs_ambiguous	1	1	0	0	0	0	2
Sisters_1	0	0	0	0	0	0	0
Sisters_2	0	0	0	0	0	0	0
Sisters_3	0	0	0	0	0	0	0
Hemizygous X	0	0	0	0	0	3	3
Triploid	0	0	1	1	0	0	2
ambiguous	1	1	1	1	1	1	6
total	11	11	11	11	11	11	66

Sterile F1

Chromosome genotype	Chr. I	Chr. II	Chr. III	Chr. IV	Chr. V	Chr. X	total
Homologs_1	2	2	3	1	1	2	11
Homologs_2	3	3	0	2	2	3	13
Homologs_3	3	2	4	4	4	1	18
Homologs_ambiguous	0	1	0	0	0	0	1
Sisters_1	0	0	1	0	0	0	1
Sisters_2	0	0	0	0	0	0	0
Sisters_3	0	0	0	0	1	0	1
Hemizygous X	0	0	0	0	0	2	2
Triploid	0	0	0	1	0	0	1
ambiguous	2	2	2	2	2	2	12
total	10	10	10	10	10	10	60

Fertile and Sterile F1

Chromosome genotype	Chr. I	Chr. II	Chr. III	Chr. IV	Chr. V	Chr. X	total
Homologs_1	2	3	3	7	2	4	21
Homologs_2	7	3	1	3	4	5	23
Homologs_3	8	10	12	6	11	4	51
Homologs_ambiguous	1	2	0	0	0	0	3
Sisters_1	0	0	1	0	0	0	1
Sisters_2	0	0	0	0	0	0	0
Sisters_3	0	0	0	0	1	0	1
Hemizygous X	0	0	0	0	0	5	5
Triploid	0	0	1	2	0	0	3
ambiguous	3	3	3	3	3	3	18
total	21	21	21	21	21	21	126

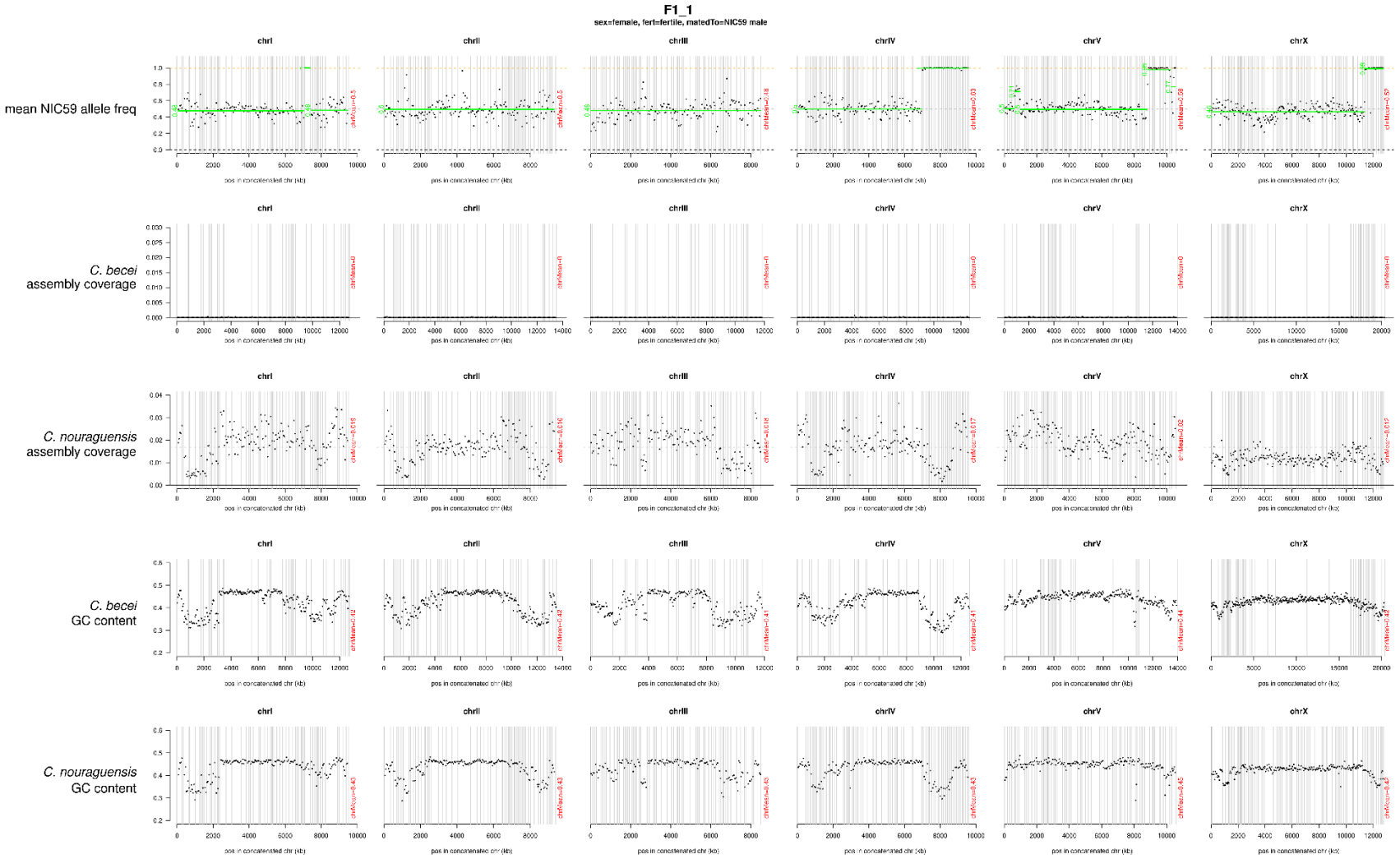
Figure 4.12. Fertile F1 inherit two randomly selected homologous chromatids from each maternal bivalent. Related to Figure 4.3 and Figure 4.4.

(A) The six possible ways of combining two of the four genetically distinct chromosomes in a bivalent are illustrated, with their expected genotypic signature below them. There are five distinct genotypic signatures. Two result from combining sister-chromatids and are called "Sisters_1" and "Sisters_2". Three result from combining homologous chromosomes and are called "Homologs_1", "Homologs_2" and "Homologs_3". **(B)** A table summarizing the genotype of each maternal chromosome for each sequenced F1 individual. The F1's sex and fertility are noted. Each individual's ploidy is inferred from the genotyping data. The genotype "Homologs_ambiguous" refers to chromosomes that are heterozygous in their centers, but one end of the chromosome is not obviously heterozygous (N/J) or homozygous for either NIC59 or JU1825. The genotype "Hemizygous X" refers to hemizygous X-chromosomes in males that have half the read depth of the autosomes. The genotype "Triploid" refers to chromosomes that have three copies instead of two based on relative read depth and whole chromosome genotype. The "Sisters_3" genotype refers to chromosomes that have inherited two non-recombinant JU1825 chromatids. Three fertile females (F1_8, F1_11 and F1_39) have NIC59 and JU1825 alleles in the center of their chromosomes, but exhibit a slight skew from the expected 0.50 NIC59 allele frequency (Figure 4.13). We hypothesize that this skew is due to contaminating DNA derived from backcrossing each female when testing her fertility. For example, a viable female that was backcrossed to a JU1825 male would carry JU1825 sperm and therefore JU1825 DNA in her spermatheca. In this case, the contaminating DNA would skew the female's entire genome to a lower NIC59 SNP frequency. Consistent with this, in all three females the more abundant allele matches the genotype of the male she was backcrossed to (Table 4.1). Correcting for this potential backcrossing contamination, the genotypes of all the chromosomes in these three females are consistent with the inheritance of two homologous chromatids. One fertile

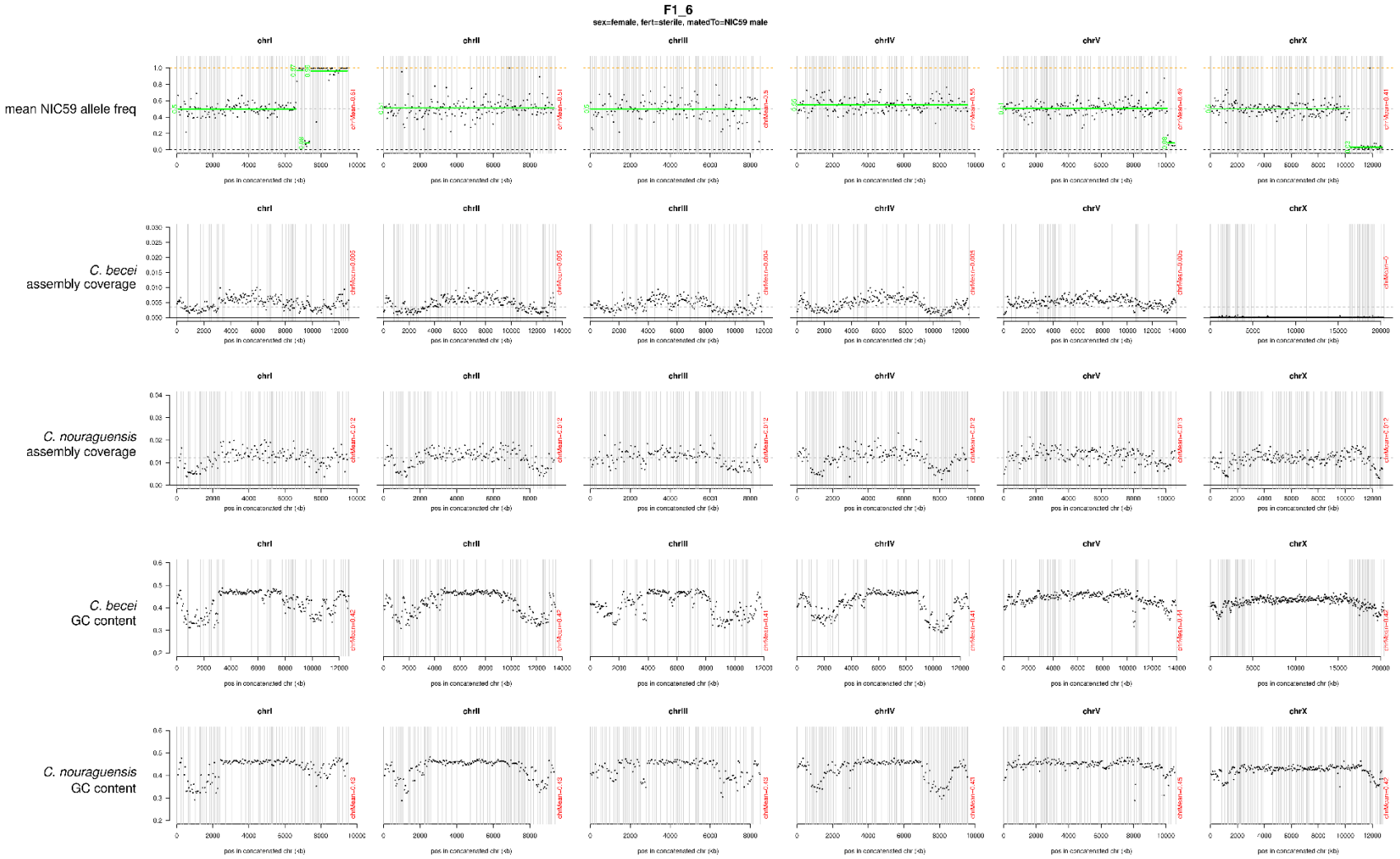
female (F1_25) and two sterile females (F1_18 and F1_26) have a very low proportion of NIC59 reads across their entire genome (Figure 4.13). We hypothesize this skew is due to contaminating DNA from JU1825 males during fertility testing and poor lysis of the female (Table 4.1). The low proportion of NIC59 reads made it difficult to infer actual maternal chromosome genotypes and are therefore categorized as "ambiguous". **(C)** Tables summarizing the frequency of genotypes per chromosome in all fertile F1, all sterile F1, or a combination of all fertile and sterile F1.

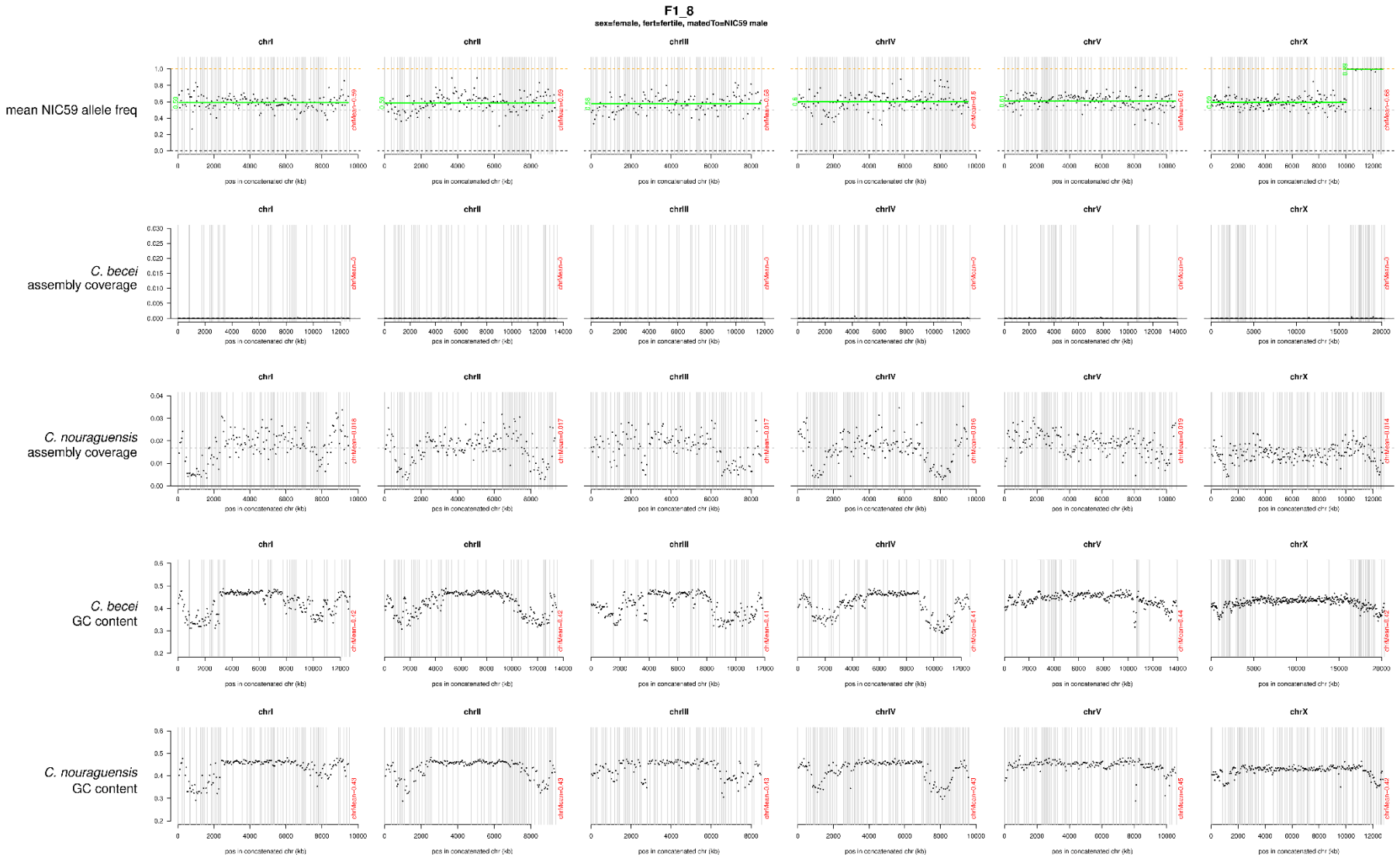
Figure 4.13. Fertile F1 inherit two randomly selected homologous chromatids from each maternal bivalent. Related to Figure 4.3 and Figure 4.4.

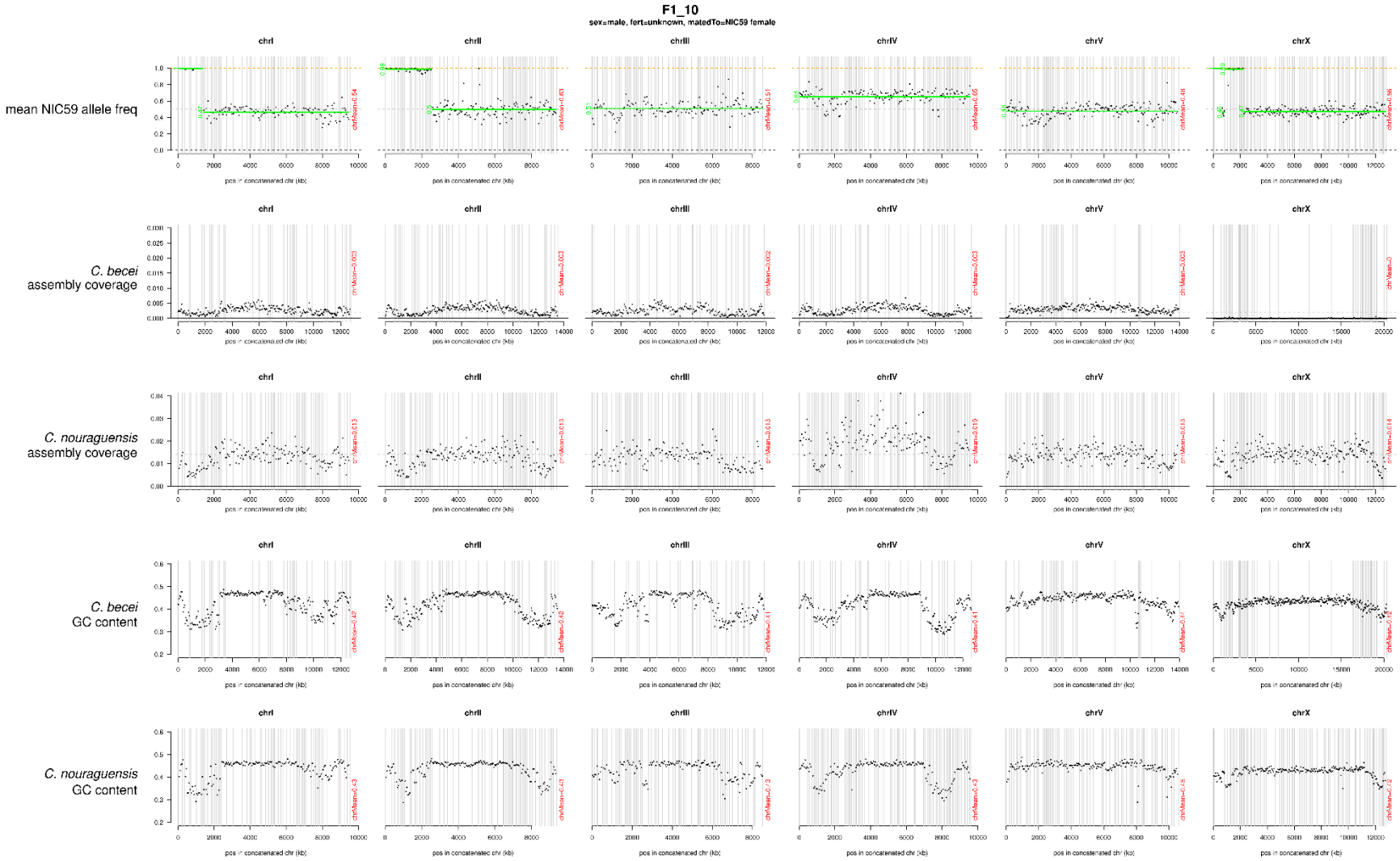
Each of the following pages contains plots describing whole-genome sequencing data of either a rare viable F1 individual or a control DNA sample. The sample name is at the top of each page, along with the individual's sex, fertility and strain it was backcrossed to for fertility testing (if applicable). Each page has five rows of plots. The first row shows the genotypes of the sample's *C. nouraguensis* maternal chromosomes (i.e. average NIC59 allele frequency in 50-kb windows across the *C. nouraguensis* assembly). Haplotype change points and average allele frequency for each segment are shown by the green horizontal lines. The second and third rows show the sample's average read coverage of the *C. becei* and *C. nouraguensis* assemblies in 50-kb windows. The fourth and fifth rows show the average GC content of the *C. becei* and *C. nouraguensis* assemblies in 50-kb windows. The gray vertical lines represent breaks between scaffolds.

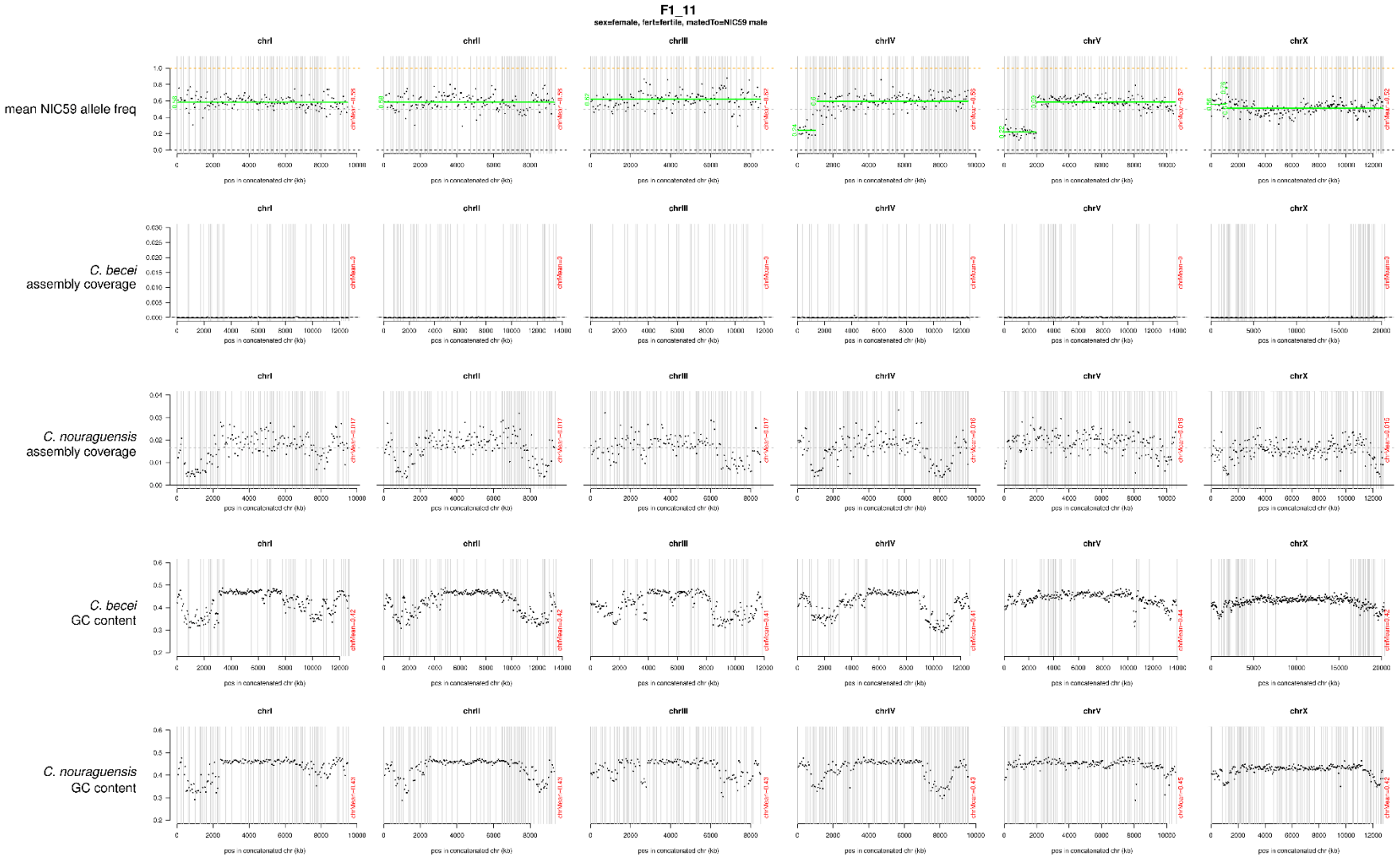

F1_1
sex=female, fert=fertile, matedTo=NIC59 male
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

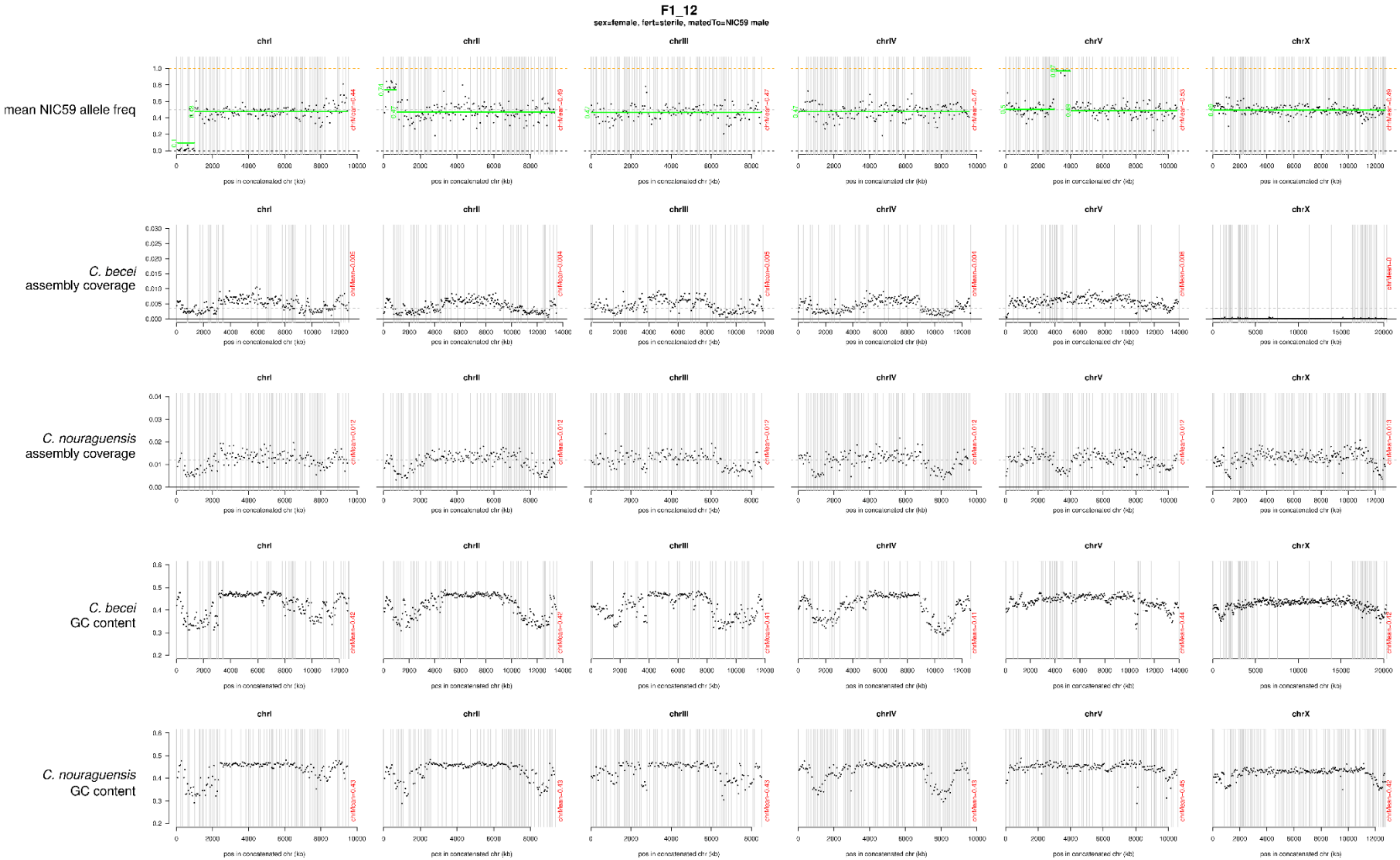
F1_5
sex=female, fert=fertile, mate=To=NIC59 male
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

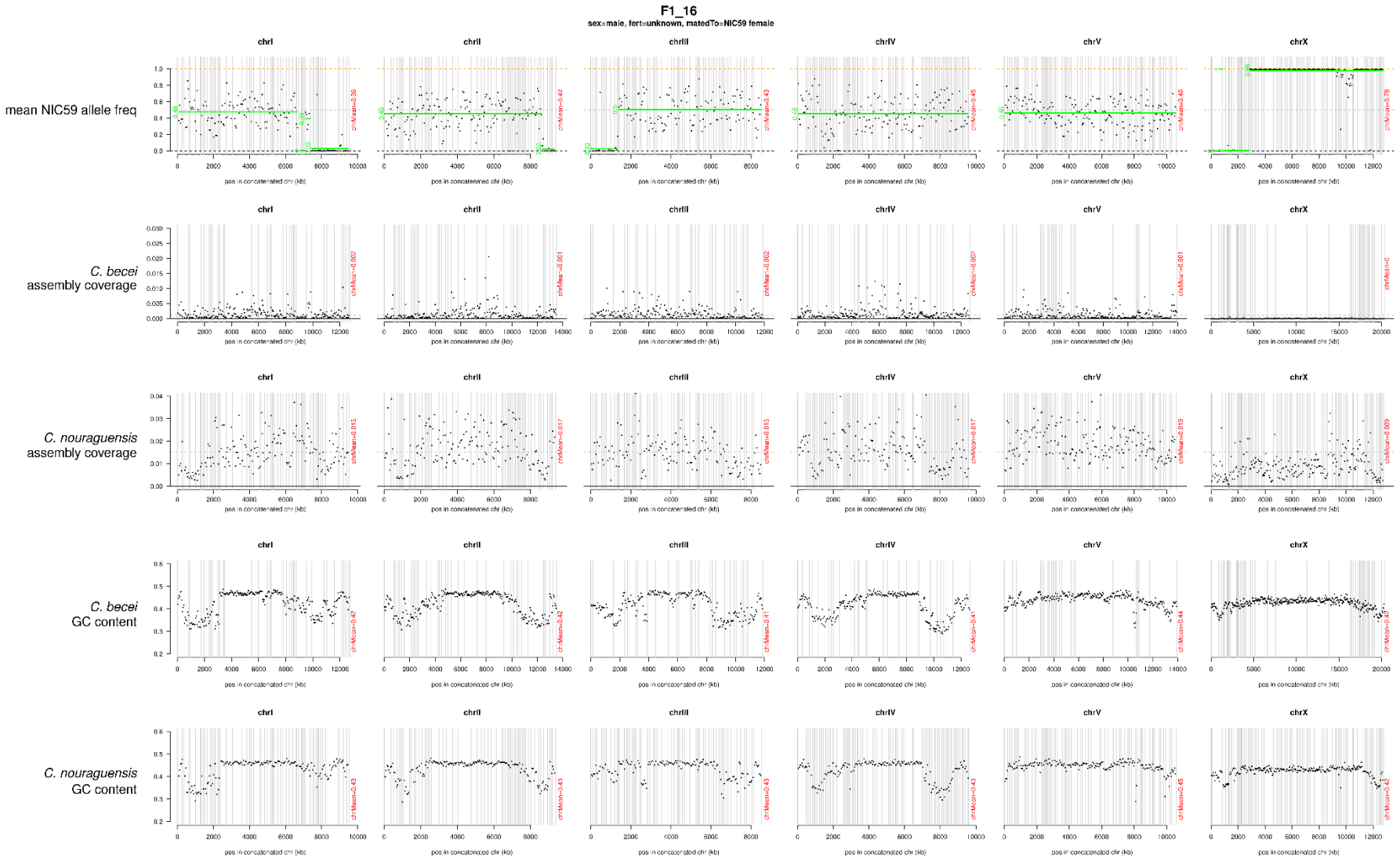

F1_6
sex=female, fert=sterile, matedTo=NIC59 male
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

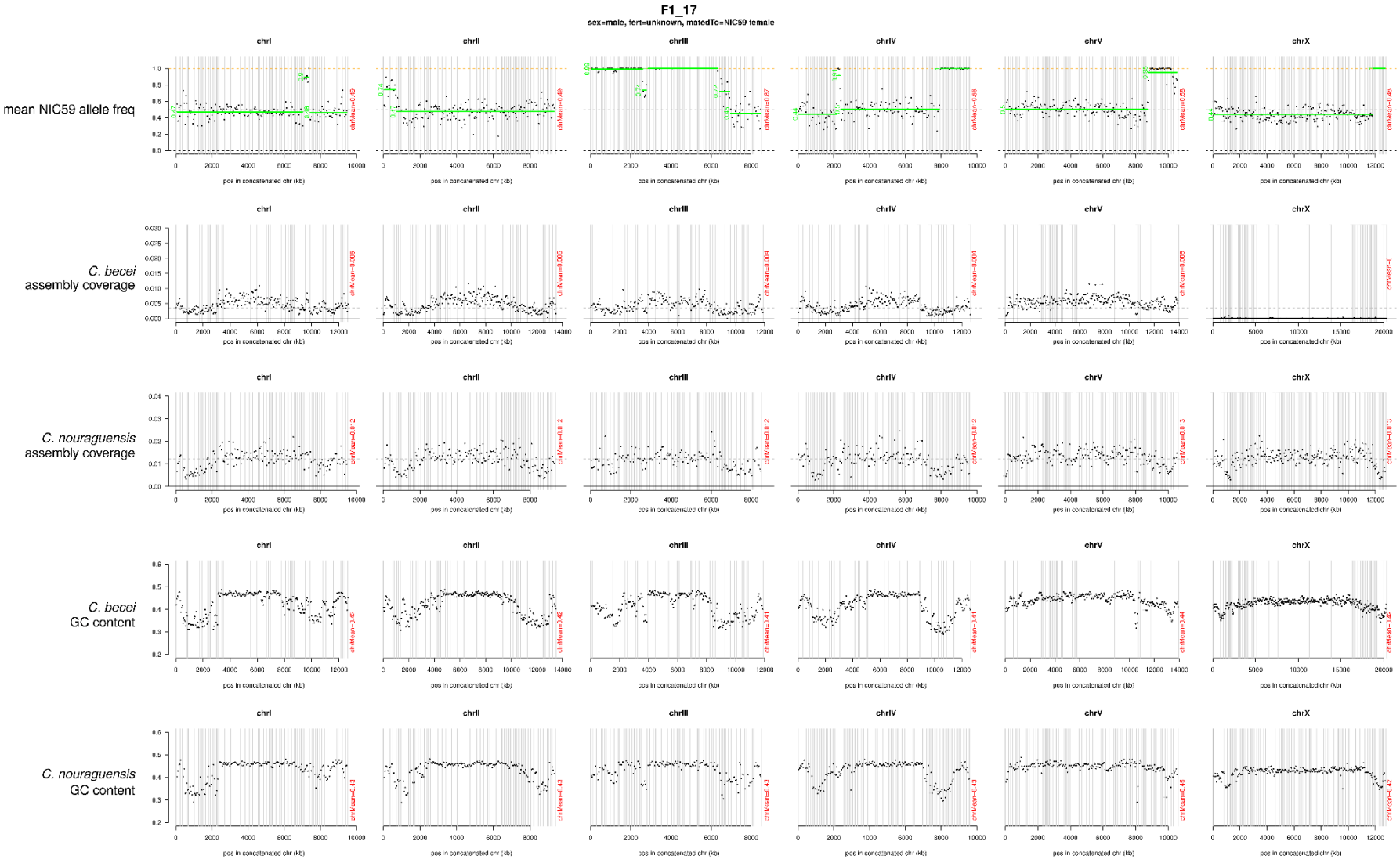

F1_8
sex=female, fert=fertile, matedTo=NIC59 male
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

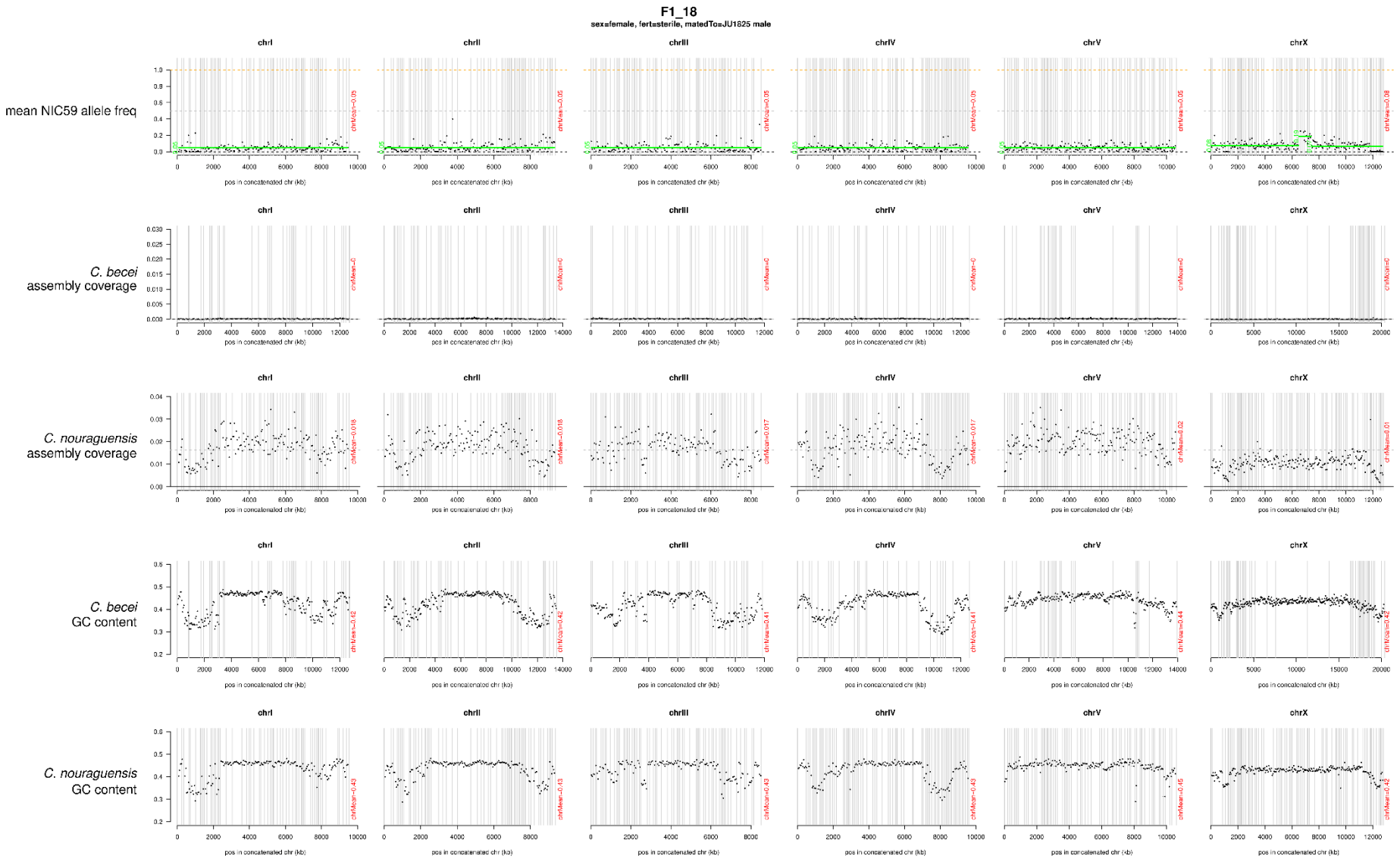

F1_10
sex=male, fert=unknown, matedTo=NIC59 female
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

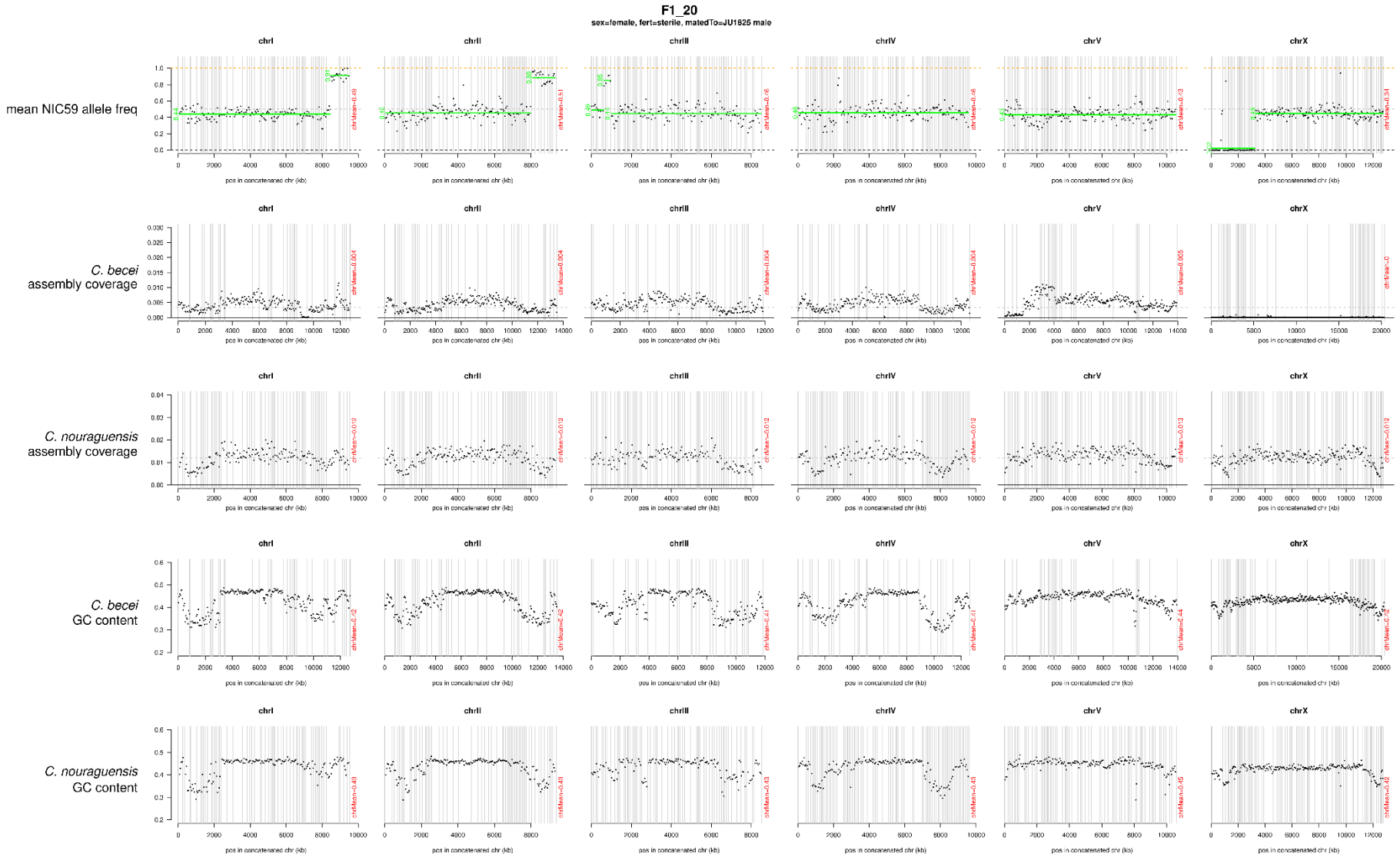

F1_11
sex=female, fert=fertile, matedTo=NIC59 male
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

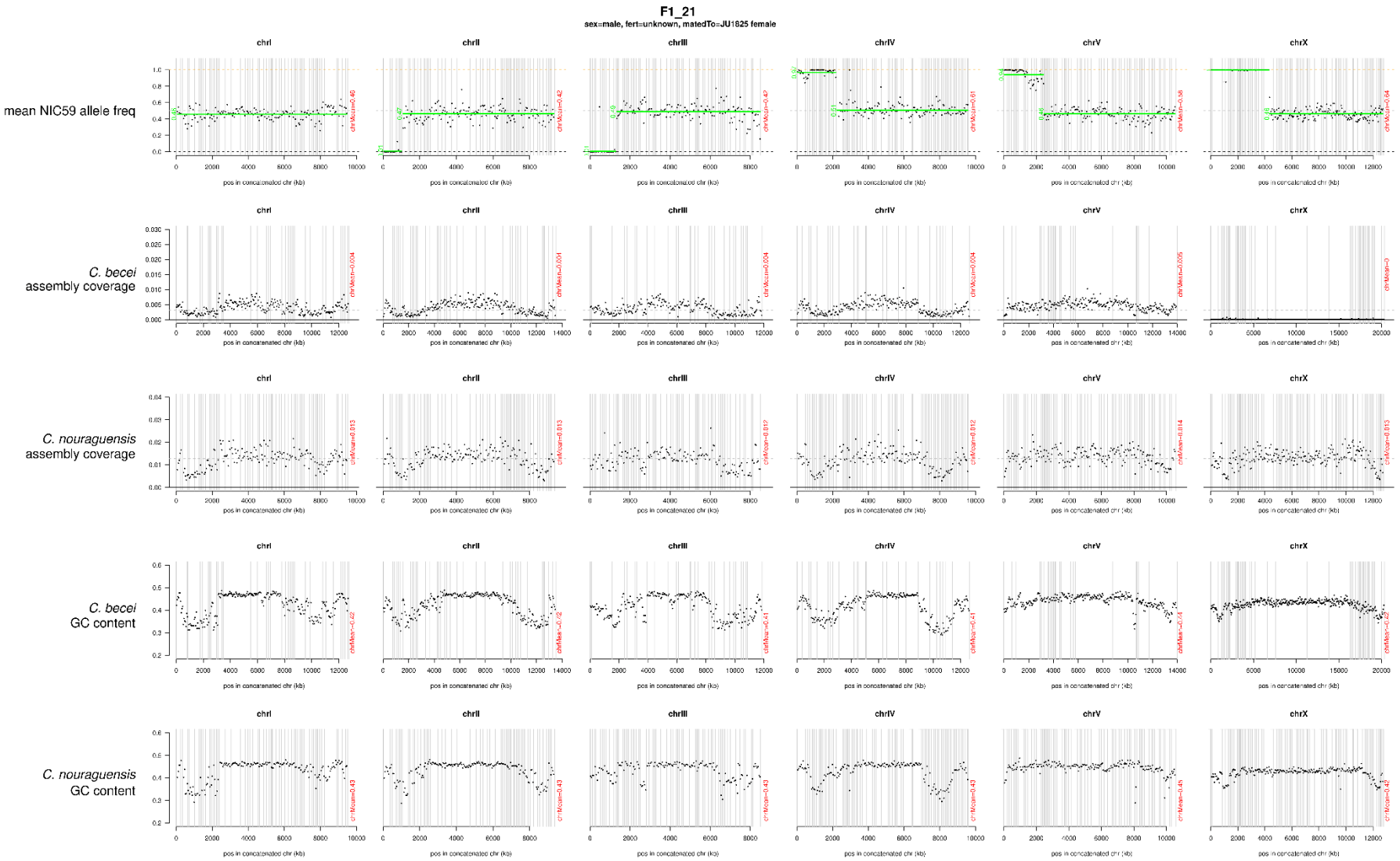

F1_12
sex=female, fert=sterile, matedTo=NIC59 male
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

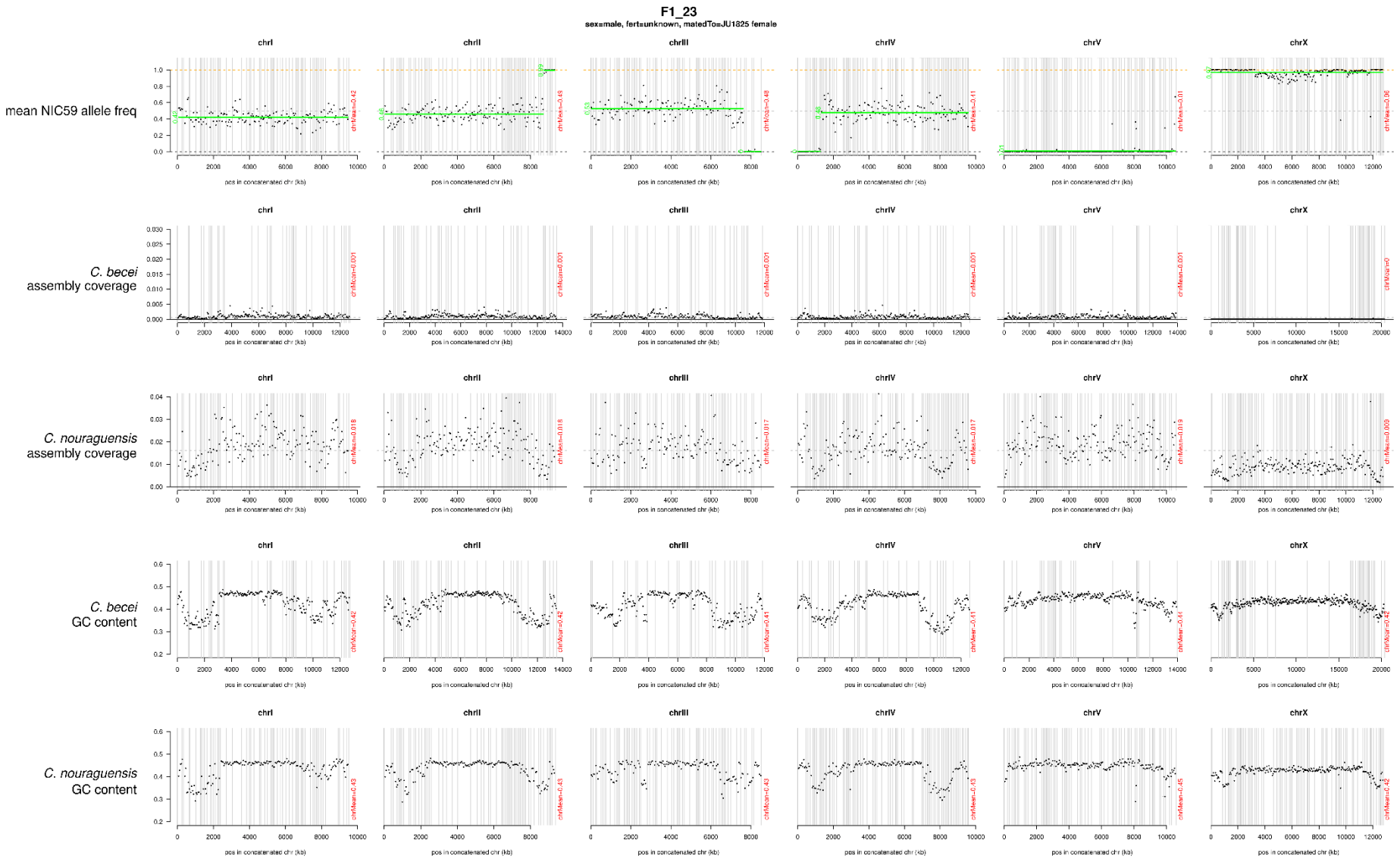

F1_16
sex=male, fert=unknown, matedTo=NIC59 female
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

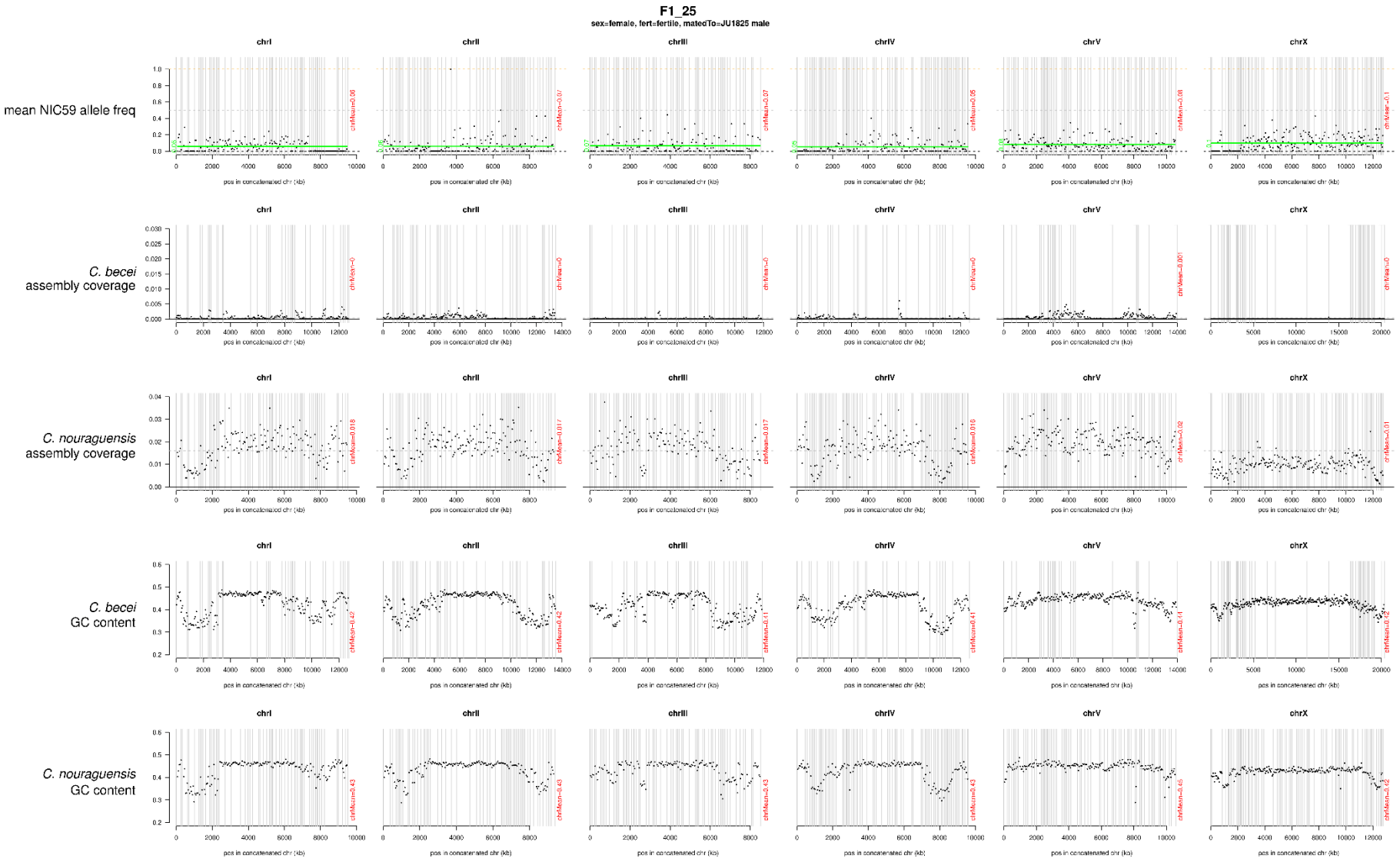

F1_17
sex=male, fert=unknown, matedTo=NIC59 female
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

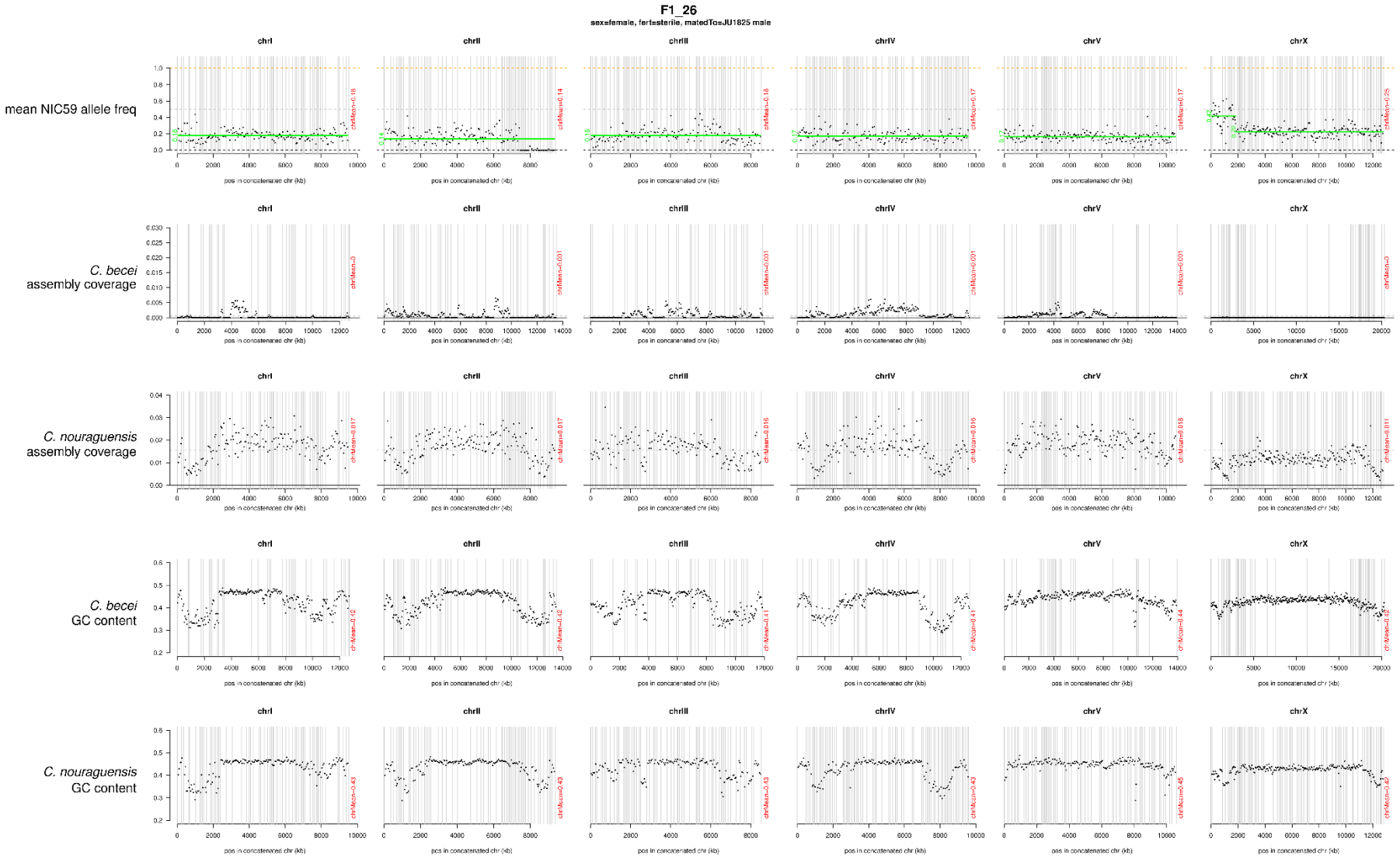

F1_18
sex=female, fert=sterile, matedTo=JU1825 male
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

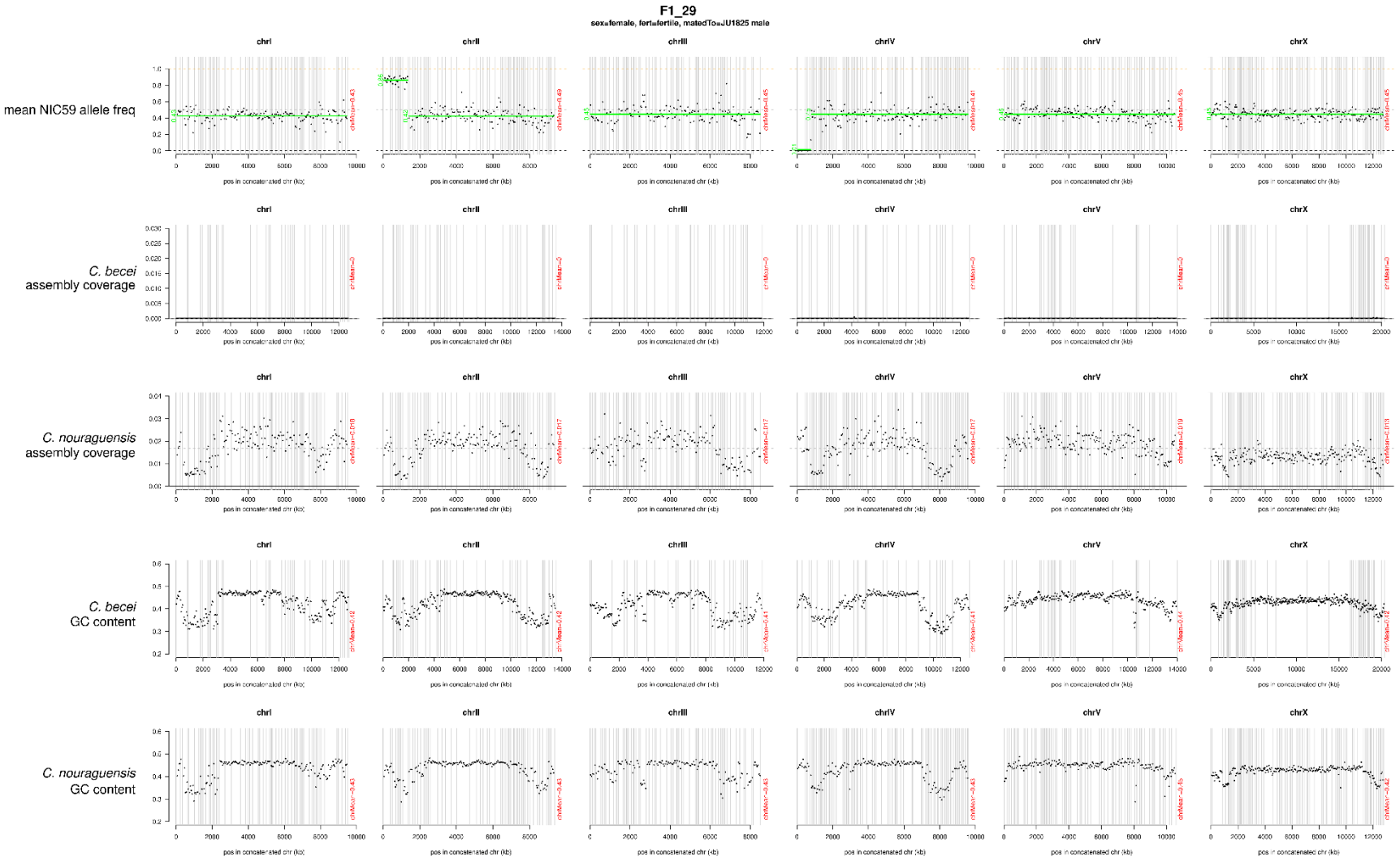

F1_20
sex=female, fert=sterile, matedTo=JU1825 male
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

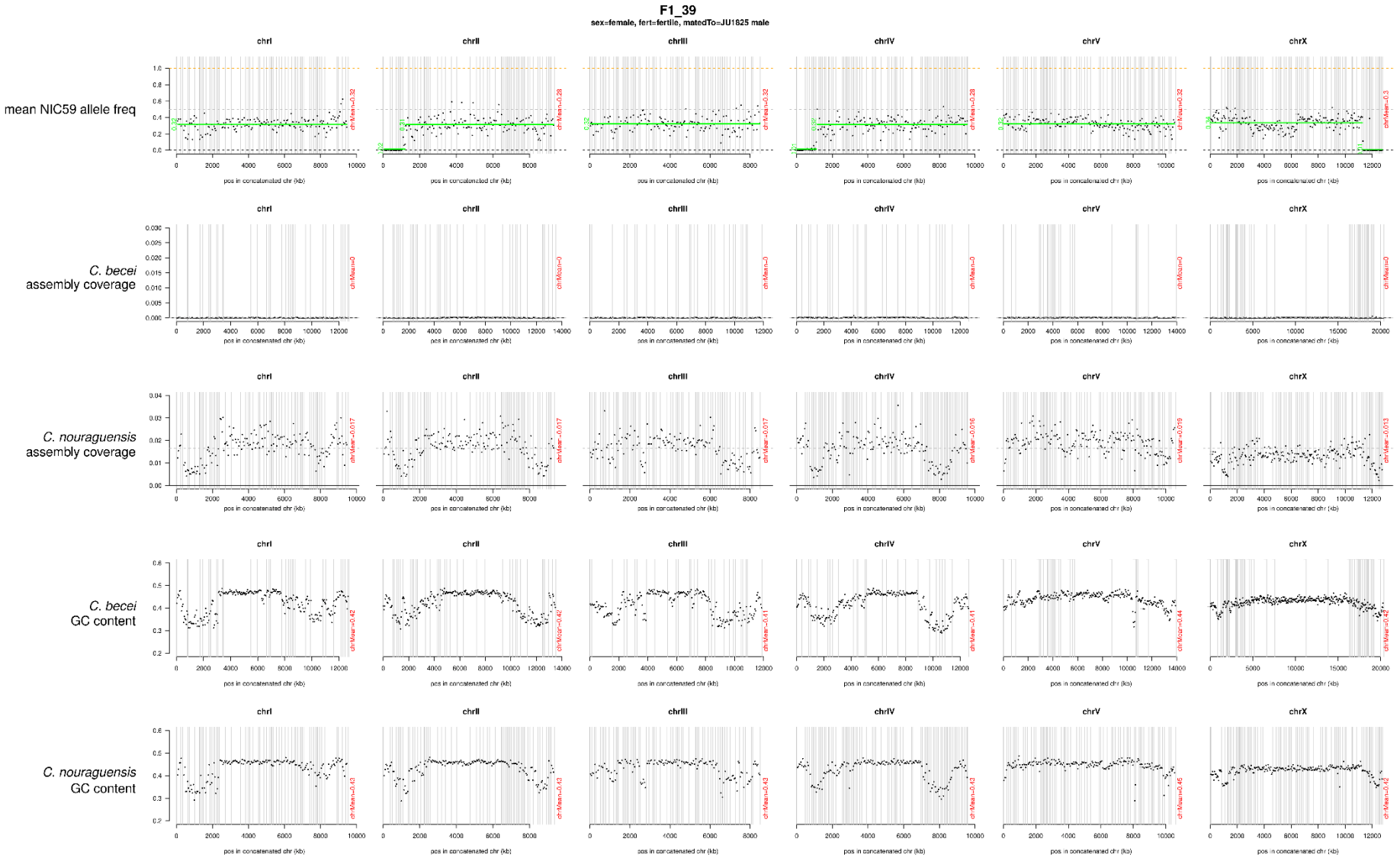

F1_21
sex=male, fert=unknown, matedTo=JU1825 female
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

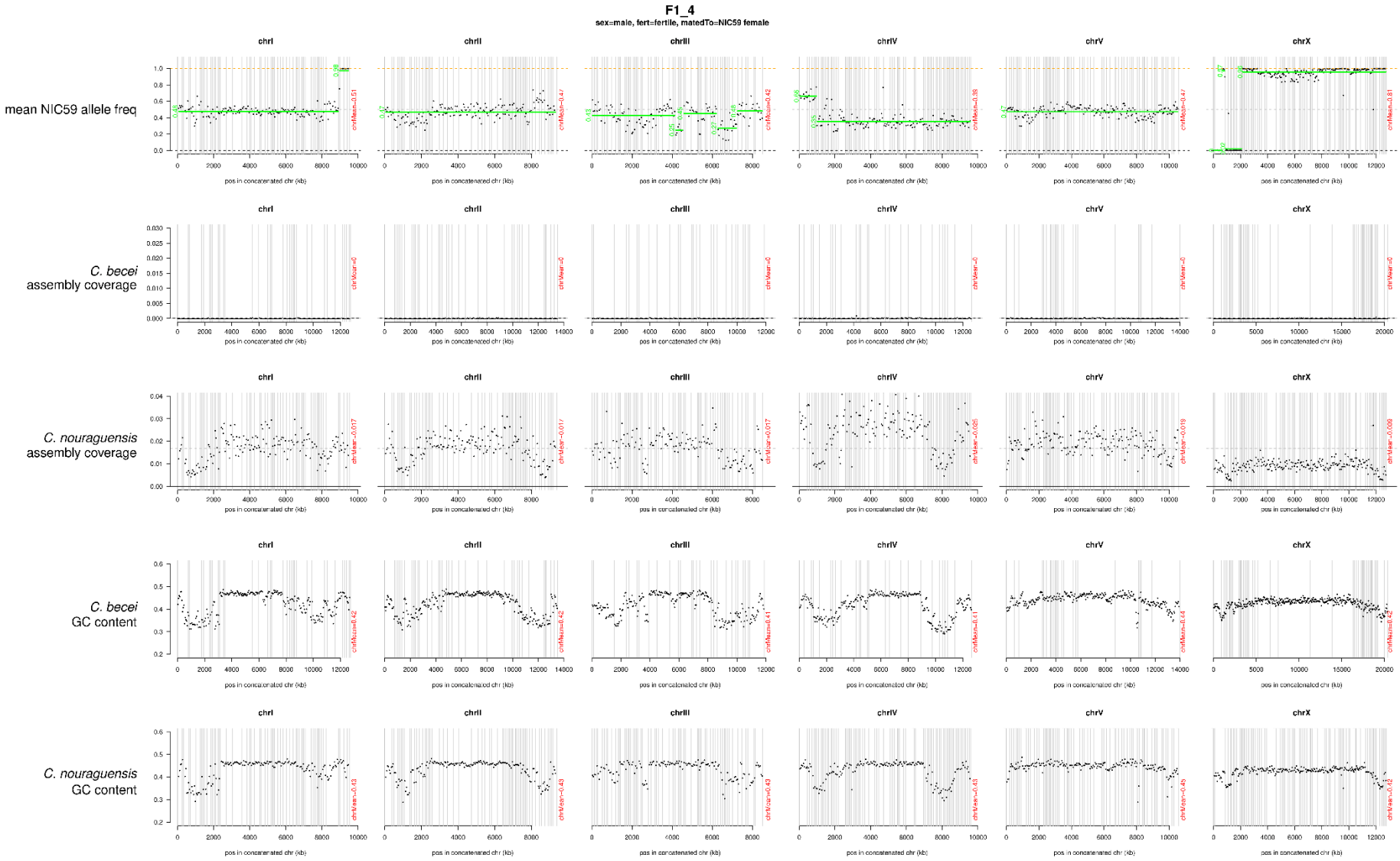

F1_23
sex=male, fert=unknown, matedTo=JU1825 female
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

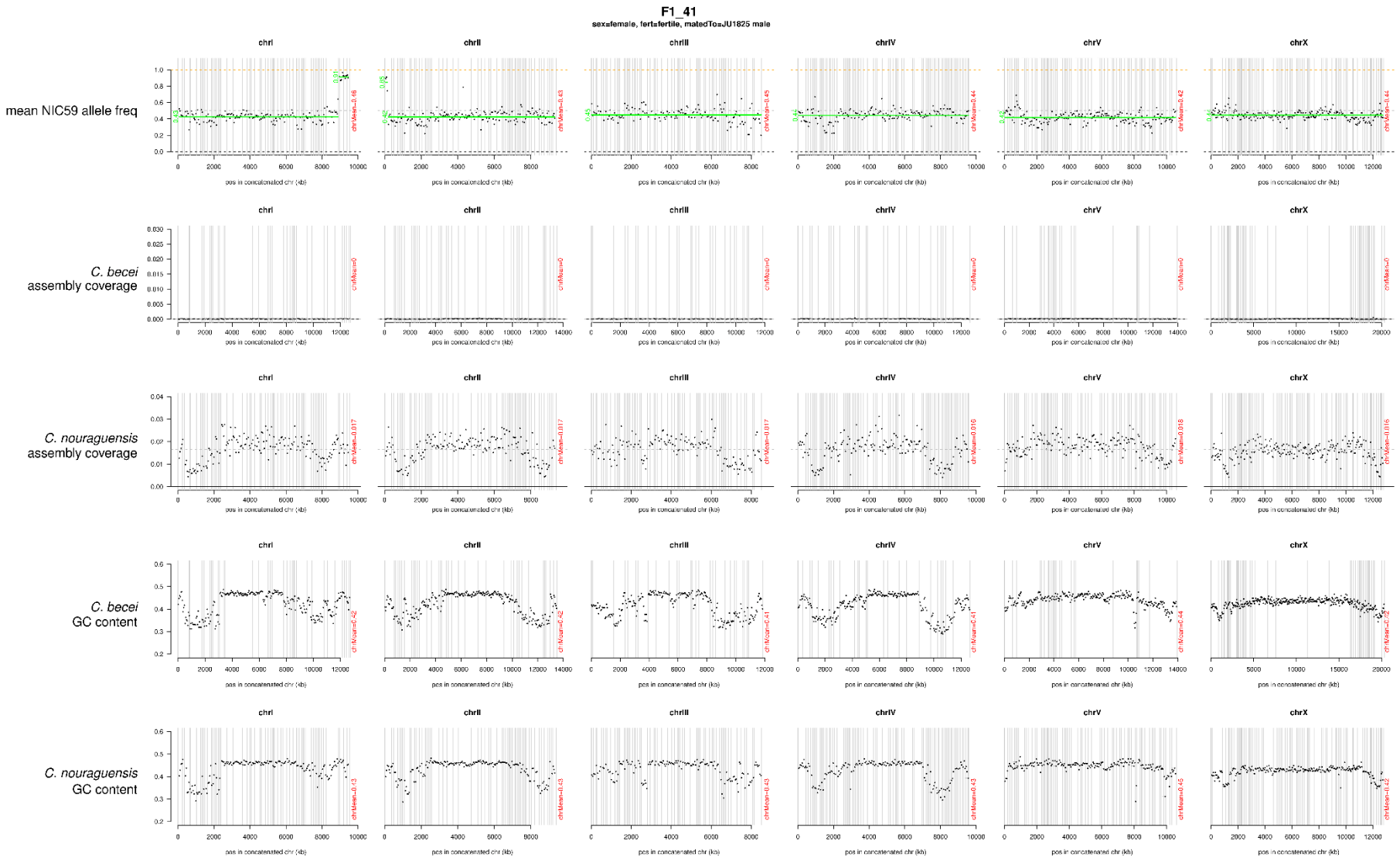

F1_25
sex=female, fert=fertile, matedTo=JU1825 male
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

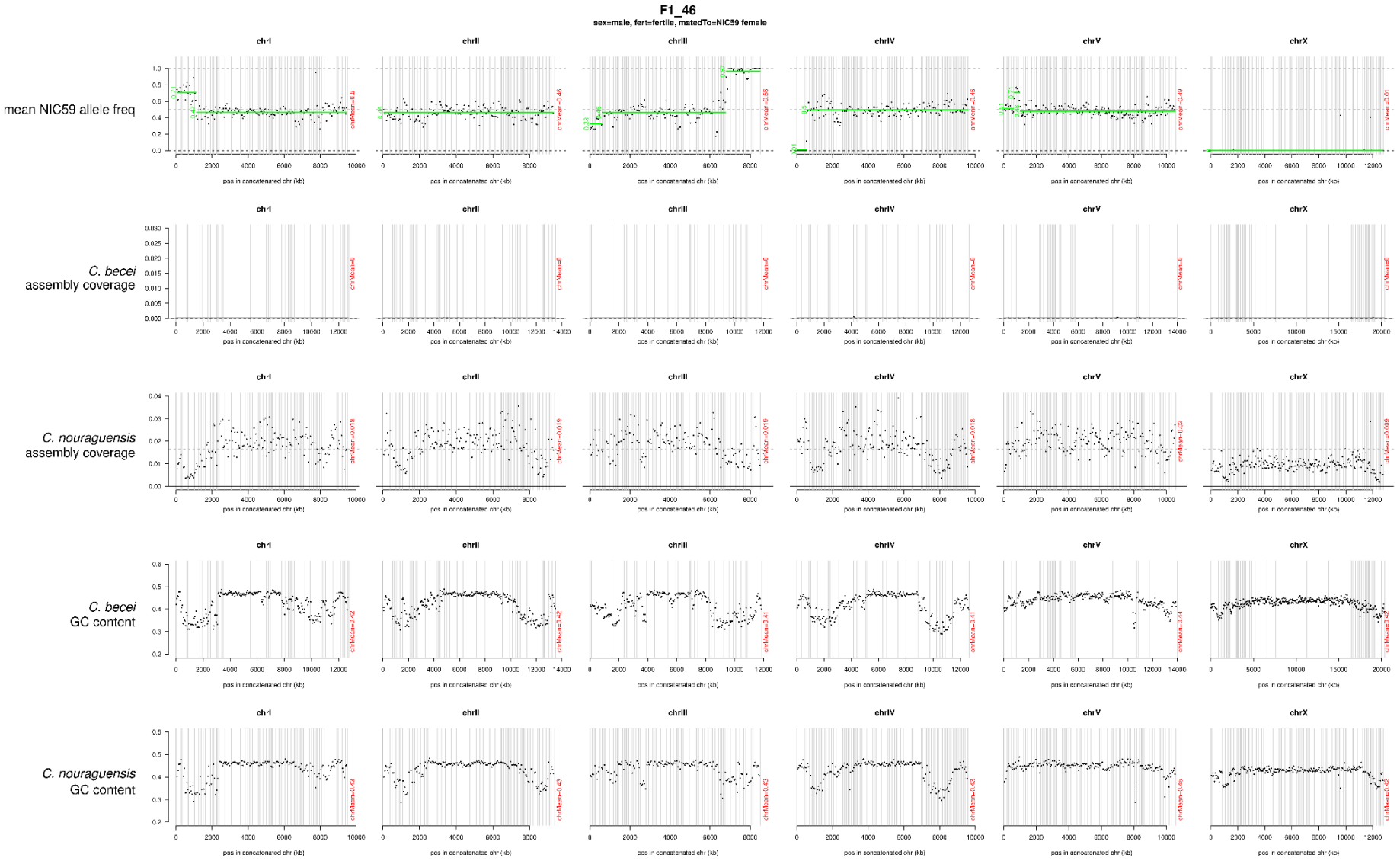

F1_26
sex=female, fert=sterile, matedTo=JU1825 male
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

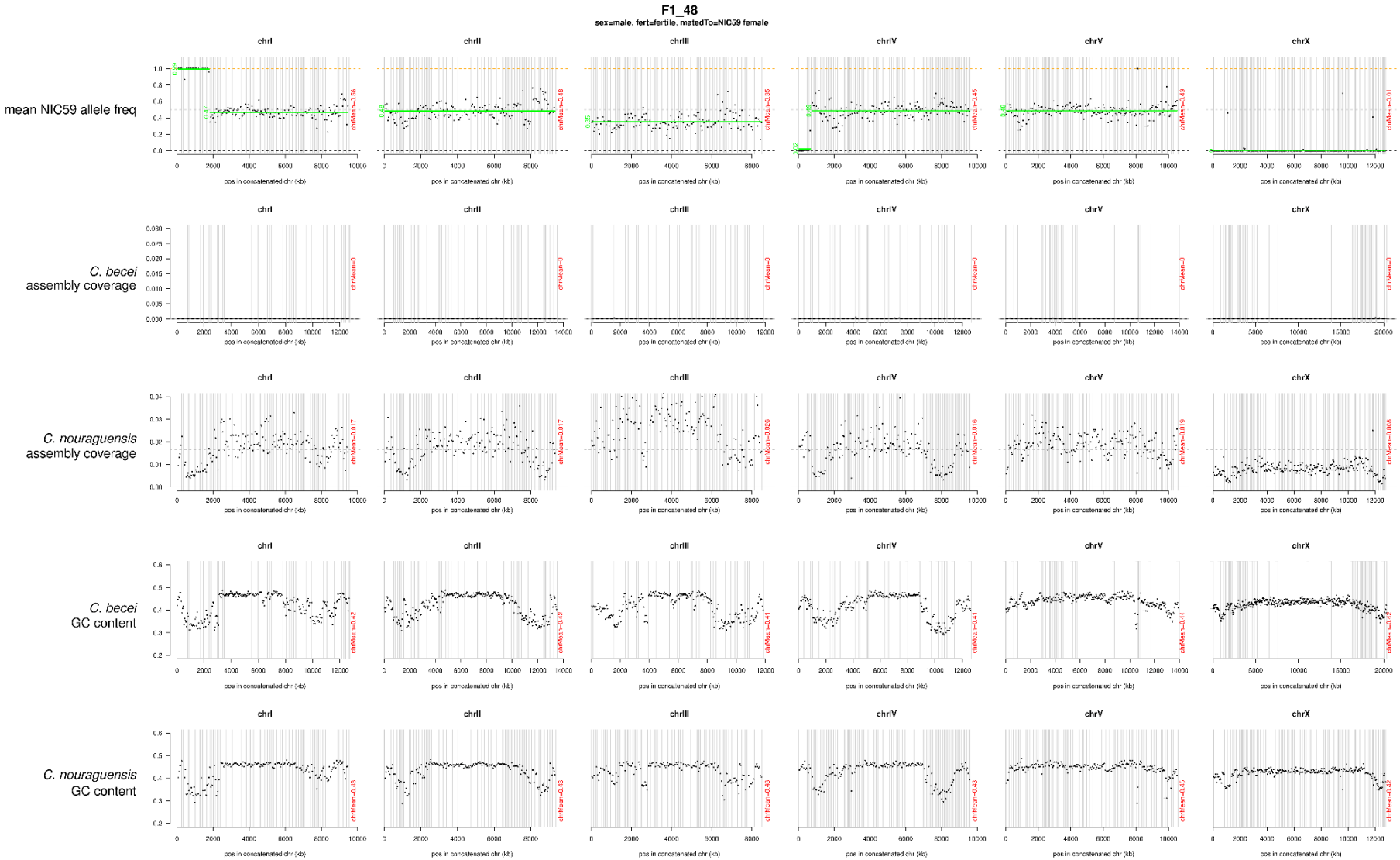

F1_29
sex=female, fert=fertile, matedTo=JU1825 male
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

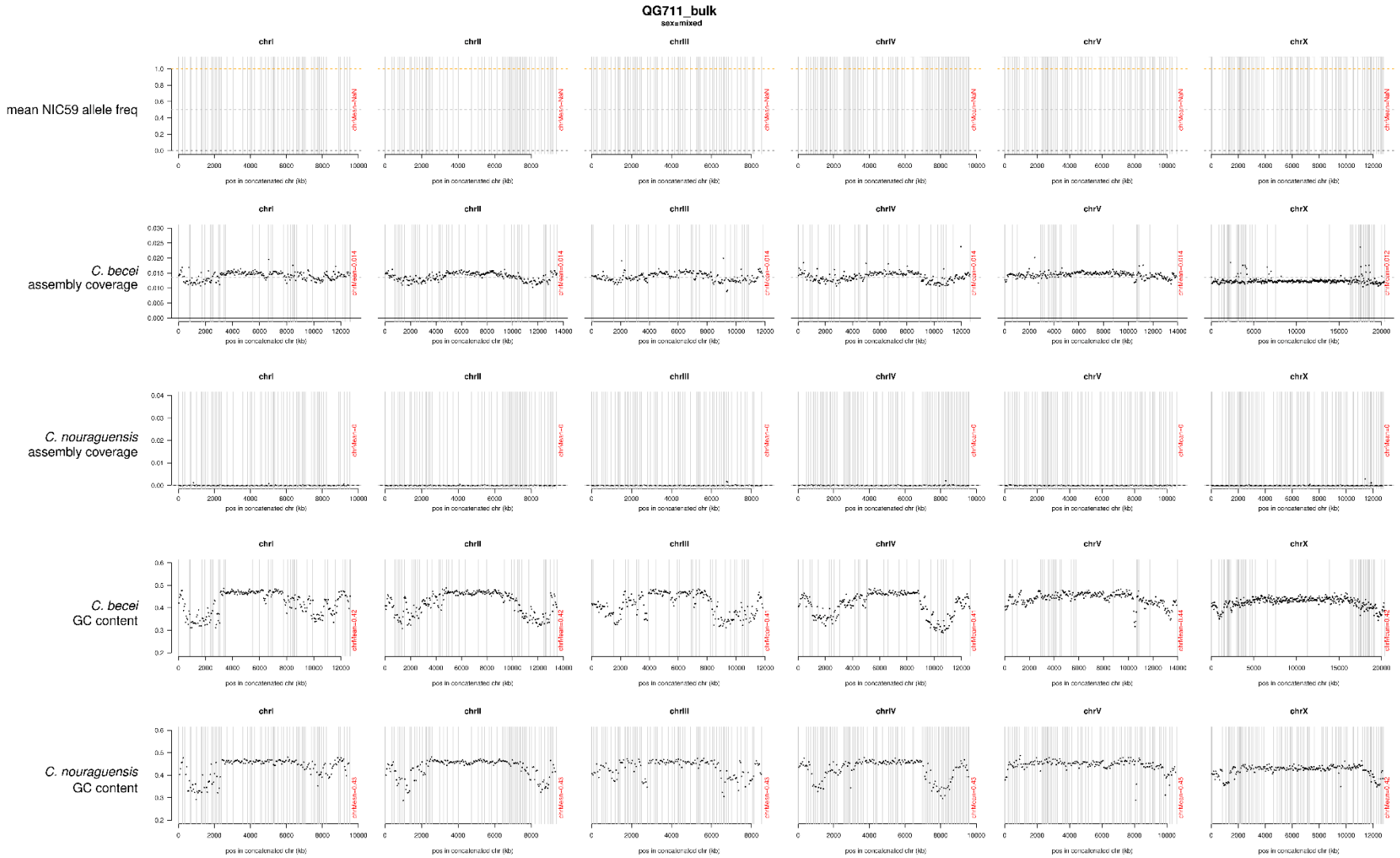

F1_39
sex=female, fert=fertile, matedTo=JU1825 male
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

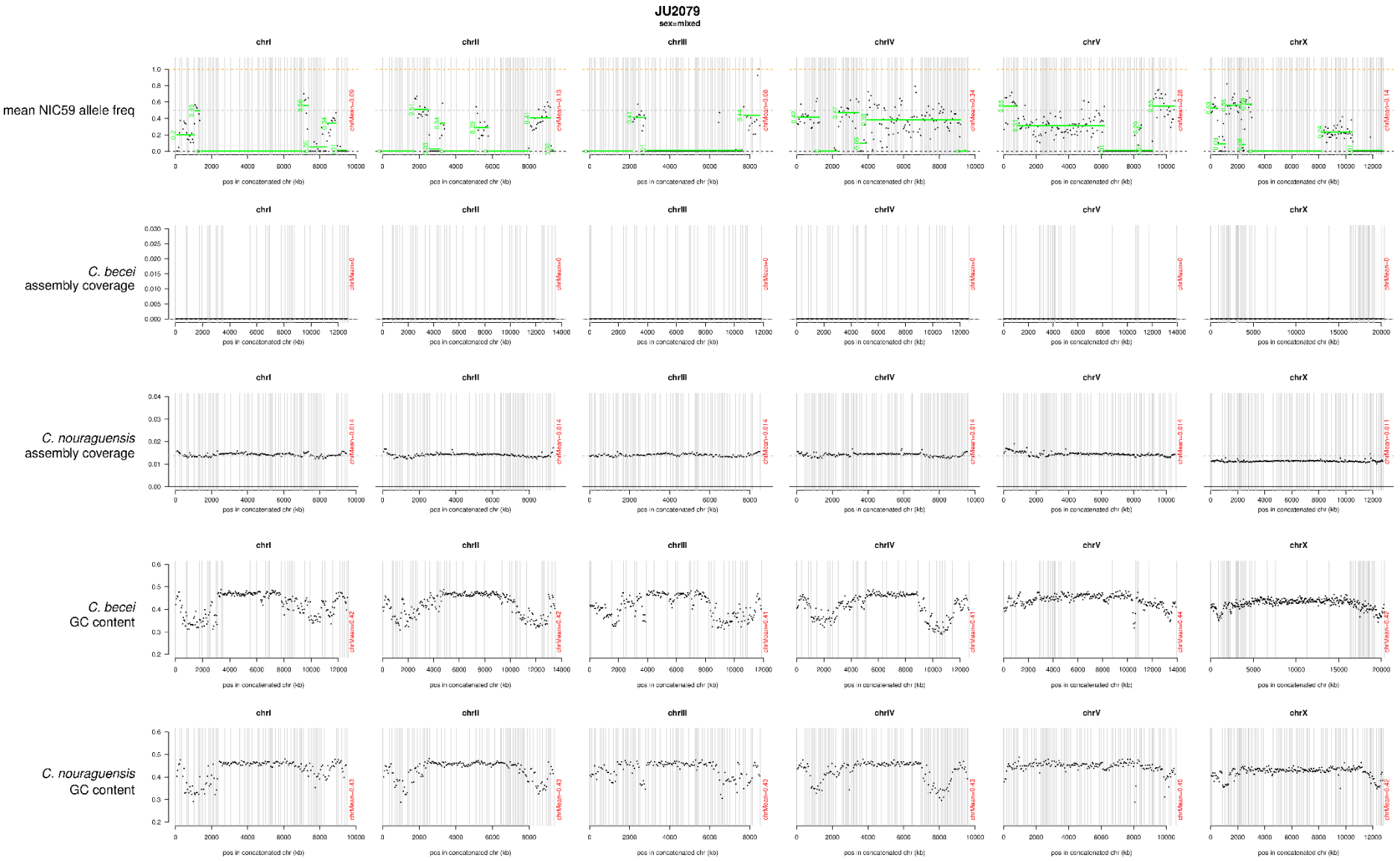

F1_4
sex=male, fert=fertile, matedTo=NIC59 female
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

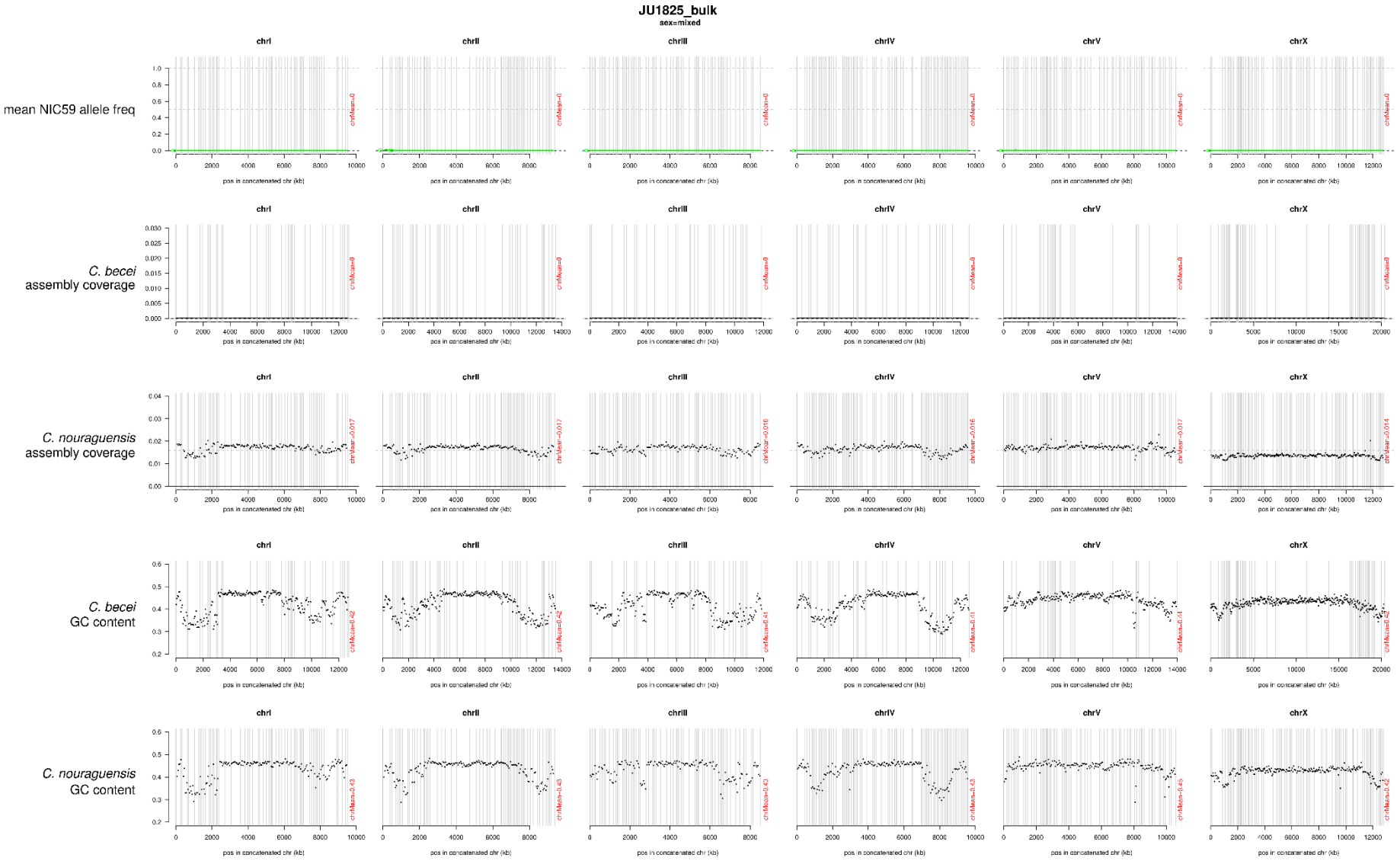

F1_41
sex=female, fert=fertile, matedTo=JU1825 male
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

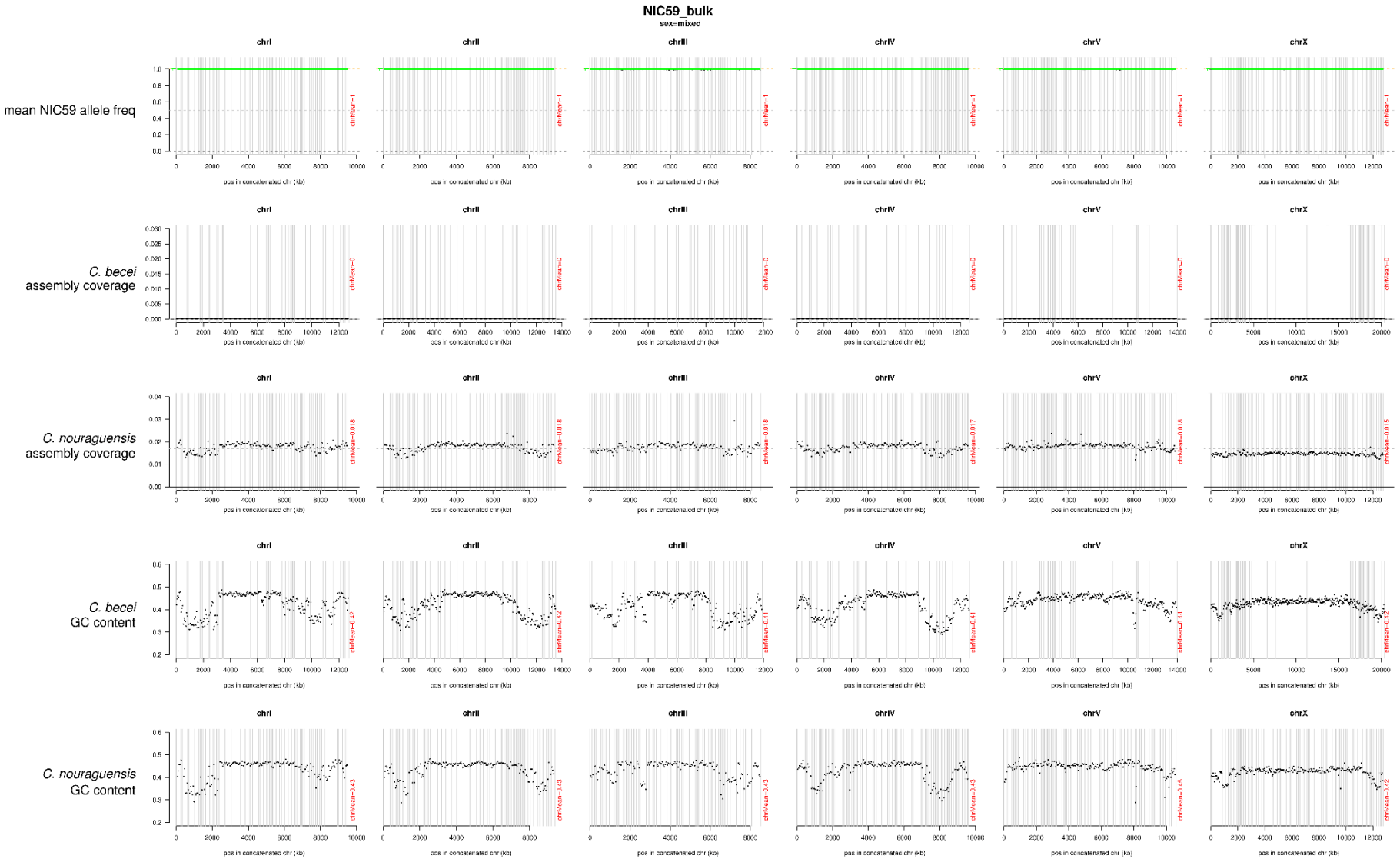

F1_46
sex=male, fert=fertile, matedTo=NIC59 female
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

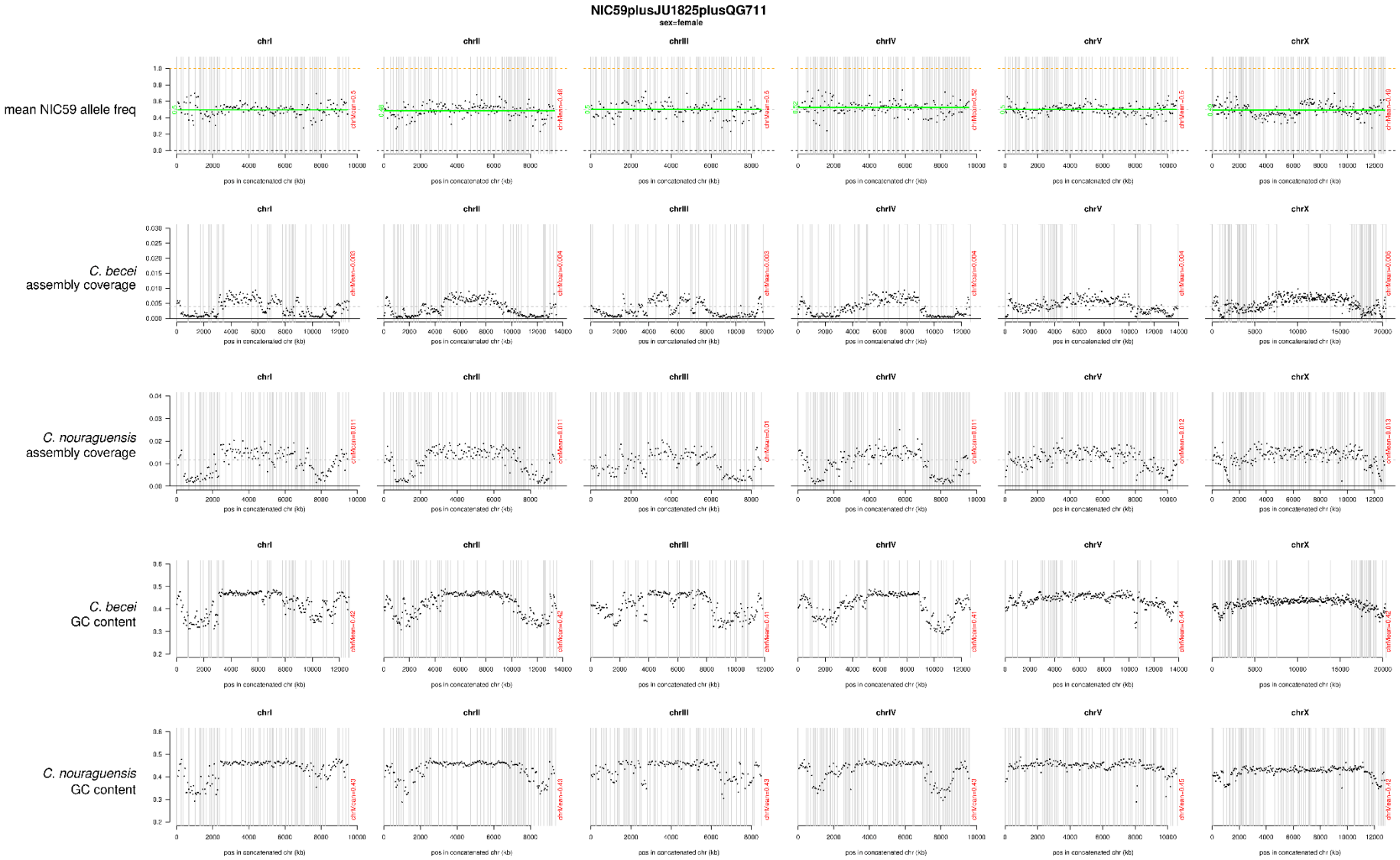

F1_48
sex=male, fert=fertile, matedTo=NIC59 female
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

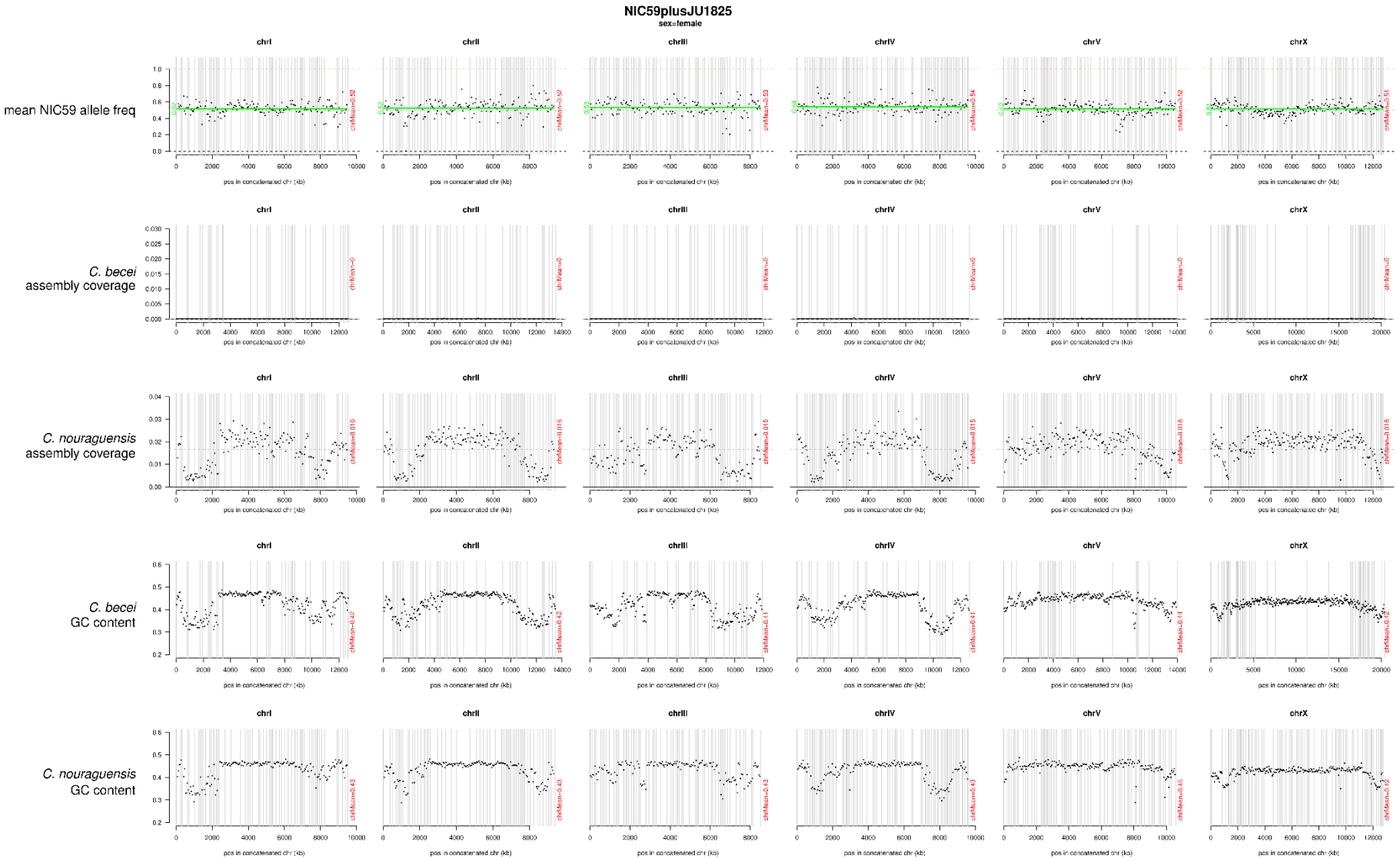

QG711_bulk
sex=mixed
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

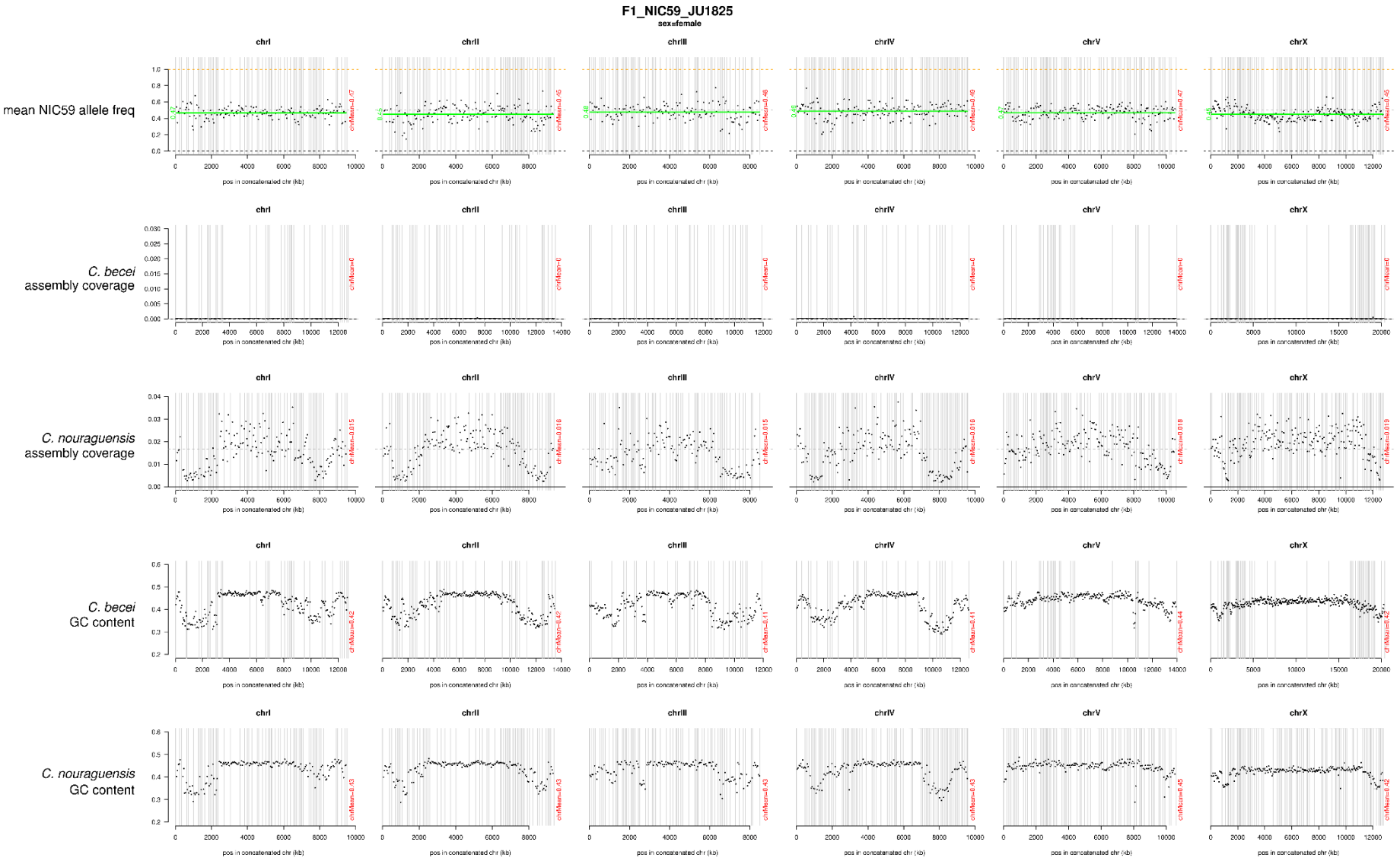

JU2079
sex=mixed
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)


JU1825_bulk
sex=mixed
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)


NIC59_bulk
sex=mixed
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)


NIC59plusJU1825plusQG711
sex=female
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)


NIC59plusJU1825
sex=female
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)


F1_NIC59_JU1825
sex=female
chrI
chrII
chrIII
chrIV
chrV
chrX
mean NIC59 allele freq
C. becei assembly coverage
C. nouraguensis assembly coverage
C. becei GC content
C. nouraguensis GC content
pos in concatenated chr (kb)

4.7 TABLES

Table 4.3. **Characteristics of the samples submitted for whole-genome sequencing. Related to Figure 3 and Figure 4.**

Whole-genome amplified F1 individuals derived from the interspecies cross are labeled as "F1_" followed by a barcode number. Sample "F1_NIC59_JU1825" is one heterozygous NIC59/JU1825 female that was whole-genome amplified and serves as a control for genome-wide heterozygosity. Sample "NIC59plusJU1825" comprises one NIC59 and one JU1825 female placed in the same tube and whole-genome amplified. This also serves as a control for genome-wide heterozygosity. Sample "NIC59plusJU1825plusQG711" comprises one NIC59, one JU1825 and one QG711 female placed into the same tube and whole-genome amplified. This serves as a triploid (diploid *C. nouraguensis* and haploid *C. becei*) control. The "NIC59_bulk", "JU1825_bulk" and "QG711_bulk" samples are genomic preps made from large populations of each strain containing mixed developmental stages and sexes: these libraries were prepared without whole-genome amplification. The reads from the "NIC59_bulk" and "JU1825_bulk" samples were used to identify fixed SNPs between the two strains. The second column shows the interspecies cross that produced each viable F1 animal. In each case, heterozygous N/J *C. nouraguensis* females were crossed to *C. becei* QG711 males. However, the heterozygous *C. nouraguensis* females could have either a NIC59 or JU1825 mitochondrial genotype (denoted in parentheses). A fraction of each sample's reads (Illumina 50-bp paired end) derive from *E.coli* (their food). The *E. coli* fraction is quite high in some samples, likely due to inefficient cleaning of single worms (see Materials and Methods). Of the reads that map to *Caenorhabditis* nuclear genomes, the approximate coverage and percent of all worm reads that map to each *Caenorhabditis* nuclear genome are shown.

						After filtering, number of reads matching						After filtering, percent of all assigned reads matching			After filtering, percent of all worm nuclear genome reads matching		Approximate coverage	
Sample name	Cross F1 derived from	F1 sex	F1 crossed to	F1 fertility	Total # reads	*E. coli*	Firmicutes	*C. nouraguensis* nuclear genome	*C. becei* nuclear genome	*C. nouraguensis* mitochondrial genome	*C. becei* mitochondrial genome	*E. coli*	*C. nouraguensis* nuclear genome	C. becei nuclear genome	*C. nouraguensis*	*C. becei*	*C. nouraguensis* nuclear genome	*C. becei* nuclear genome
F1_1	(N); N/J F1 female x QG711 male Plate 1	female	NIC59 male	fertile	18,405,750	11,551,300	19	3,529,806	10,854	9,412	103	76.5	23.4	0.1	99.7	0.3	2.4	0.0
F1_5	(N); N/J F1 female x QG711 male Plate 2	female	NIC59 male	fertile	4,709,920	306,730	6	3,291,061	6,828	5,605	61	8.5	91.2	0.2	99.8	0.2	2.2	0.0
F1_8	(N); N/J F1 female x QG711 male Plate 2	female	NIC59 male	fertile	6,219,960	2,499,294	4	2,528,892	5,174	4,160	48	49.6	50.2	0.1	99.8	0.2	1.7	0.0
F1_11	(N); N/J F1 female x QG711 male Plate 3	female	NIC59 male	fertile	5,556,004	1,228,558	3	3,169,866	6,505	1,879	20	27.9	71.9	0.1	99.8	0.2	2.2	0.0
F1_25	(J); N/J F1 female x QG711 male Plate 3	female	JU1825 male	fertile	6,489,980	5,089,384	3	369,912	8,714	499	8	93.1	6.8	0.2	97.7	2.3	0.3	0.0
F1_29	(J); N/J F1 female x QG711 male Plate 4	female	JU1825 male	fertile	7,097,772	667,344	13	4,884,750	9,066	1,495	16	12.0	87.8	0.2	99.8	0.2	3.3	0.0
F1_39	(J); N/J F1 female x QG711 male Plate 1	female	JU1825 male	fertile	4,858,964	688,291	5	3,186,443	7,190	3,346	14	17.7	82.0	0.2	99.8	0.2	2.2	0.0
F1_41	(J); N/J F1 female x QG711 male Plate 2	female	JU1825 male	fertile	18,011,610	1,112,932	20	12,756,921	21,799	2,298	19	8.0	91.8	0.2	99.8	0.2	8.7	0.0
F1_4	(N); N/J F1 female x QG711 male Plate 1	male	NIC59 female	fertile	13,552,874	151,066	6	10,195,147	22,476	5,451	51	1.5	98.3	0.2	99.8	0.2	7.0	0.0
F1_46	(N); N/J F1 female x QG711 male Plate 1	male	NIC59 female	fertile	20,496,550	4,494	15	15,463,159	35,591	7,759	83	0.0	99.7	0.2	99.8	0.2	10.6	0.0
F1_48	(N); N/J F1 female x QG711 male Plate 4	male	NIC59 female	fertile	5,484,366	13,691	3	4,151,082	9,759	661	6	0.3	99.4	0.2	99.8	0.2	2.8	0.0
F1_6	(N); N/J F1 female x QG711 male Plate 2	female	NIC59 male	sterile	4,477,900	930,700	7	1,932,322	705,081	1,202	15	26.1	54.1	19.8	73.3	26.7	1.3	0.4
F1_12	(N); N/J F1 female x QG711 male Plate 3	female	NIC59 male	sterile	7,884,868	753,139	7	3,978,335	1,483,331	2,085	19	12.1	64.0	23.9	72.8	27.2	2.7	0.8
F1_18	(J); N/J F1 female x QG711 male Plate 1	female	JU1825 male	sterile	3,468,706	1,496,149	3	1,373,037	7,652	96	0	52.0	47.7	0.3	99.5	0.6	0.9	0.0
F1_20	(J); N/J F1 female x QG711 male Plate 2	female	JU1825 male	sterile	10,834,856	1,163,152	16	5,203,891	1,822,729	4,026	23	14.2	63.5	22.2	74.1	25.9	3.6	1.0
F1_26	(J); N/J F1 female x QG711 male Plate 3	female	JU1825 male	sterile	3,594,908	891,072	3	1,923,182	92,400	690	4	30.6	66.1	3.2	95.4	4.6	1.3	0.0
F1_10	(N); N/J F1 female x QG711 male Plate 2	male	NIC59 female	sterile	13,848,208	400,238	7	8,508,436	1,589,956	1,200	9	3.8	81.0	15.1	84.3	15.7	5.8	0.9
F1_16	(N); N/J F1 female x QG711 male Plate 3	male	NIC59 female	sterile	15,639,042	288,535	2	10,408,444	1,046,790	16,142	172	2.5	88.5	8.9	90.9	9.1	7.1	0.6
F1_17	(N); N/J F1 female x QG711 male Plate 3	male	NIC59 female	sterile	6,247,186	79,293	15	3,443,948	1,272,137	322	9	1.7	71.8	26.5	73.0	27.0	2.4	0.7
F1_21	(J); N/J F1 female x QG711 male Plate 2	male	JU1825 female	sterile	9,834,614	137,887	6	5,701,336	1,741,990	480	5	1.8	75.2	23.0	76.6	23.4	3.9	0.9
F1_23	(J); N/J F1 female x QG711 male Plate 2	male	JU1825 female	sterile	12,059,396	1,570,440	6	7,254,493	382,635	11,575	96	17.0	78.7	4.2	95.0	5.0	5.0	0.2
F1_NIC59_JU1825	JU1825 female x NIC59 male	female			8,237,250	75,877	14	6,344,331	12,349	3,180	38	1.2	98.6	0.2	99.8	0.2	4.3	0.0
NIC59plusJU1825		female			7,045,822	319,354	15	5,091,335	10,238	2,668	21	5.9	93.9	0.2	99.8	0.2	3.5	0.0
NIC59plusJU1825plusQG711		female			11,160,274	275,223	7	5,895,338	2,620,526	1,864	620	3.1	67.0	29.8	69.2	30.8	4.0	1.4
NIC59_bulk		mixed			32,250,274	257,304	31	22,649,102	52,717	50,912	437	1.1	98.4	0.2	99.8	0.2	15.5	0.0
JU1825_bulk		mixed			36,926,106	357,701	394	27,761,970	44,405	58,749	279	1.3	98.4	0.2	99.8	0.2	19.0	0.0
QG711_bulk		mixed			28,828,734	235,533	37	120,477	21,138,043	17	41,249	1.1	0.6	98.2	0.6	99.4	0.1	11.4

Table 4.4. ***PCR primers***

The primers used in the PCR-RFLP polymorphism assays were chosen because they lie in conserved coding sequences that are easily amplified from a range of *Caenorhabditis* species (Kiontke *et al.* 2011). Primers 5.8S1 and KK-28S-22 were originally used in Kiontke et al (Kiontke *et al.* 2011). The primers used in the PCR indel polymorphism assays were generated after the assembly of the *C. nouraguensis* genome. Indel polymorphisms between NIC59 and JU1825 were found by manually screening *C. nouraguensis* (JU2079) scaffolds in IGV (Robinson *et al.* 2011) for regions where JU1825 had read coverage but NIC59 did not (and vice versa). Indel polymorphisms between *C. nouraguensis* and *C. becei* were manually found by aligning homologous sequences on scaffolds assigned to a chromosome of interest.

Primer name	Sequence (5'-3')	Chromosome	Strains distinguished	Type of polymorphism
5.8S1	CTGCGTTACTTACCACGAATTGCAGAC	I (ITS2)	*C. n* vs *C. b*	PCR-RFLP (Hind-III)
KK-28S-22	CACTTTCAAGCAACCCGAC	I (ITS2)	*C. n* vs *C. b*	PCR-RFLP (Hind-III)
oPL20	CTGATGAGCATCGTCCGAC	II (W02B12.9)	*C. n* vs *C. b*	PCR-RFLP (EcoRI)

oPL21	CCTAGCTTGCAATCCACACG	II (W02B12.9)	*C. n* vs *C. b*	PCR-RFLP (EcoRI)
oPL78	TCCTCCAGCATATCCGCTCG	III	JU1825 vs NIC59	indel
oPL79	CAATTGCACGGAGAGACTGTTCATC	III	JU1825 vs NIC59	indel
oPL181	CGTCAAGATCAGATGCGAGACG	I	JU1825 vs NIC59	indel
oPL182	CAGTTGAGAACTGCCTGTCAGAC	I	JU1825 vs NIC59	indel
oPL318	GTGGCTTGTAGAGCTTGTGC	V	*C. n* vs *C. b*	indel
oPL319	GATCAATGATAGACATGGCC	V	*C. n* vs *C. b*	indel
oPL320	TGGTGTCGAATTCTCAAATTCC	X	*C. n* vs *C. b*	indel
oPL321	TAGGGGCAGAATATTGAACACTG	X	*C. n* vs *C. b*	indel
oPL356	TATTTGTGTGTTTGCCAGAGG	V	JU1825 vs NIC59	indel
oPL357	TCGAATACATTTGGCTTCACG	V	JU1825 vs NIC59	indel

Chapter 5. CONCLUSIONS AND FUTURE DIRECTIONS

5.1 CYTOPLASMIC-NUCLEAR INCOMPATIBILITY BETWEEN WILD-ISOLATES OF *C. NOURAGUENSIS*

My first project focuses on understanding the genetic basis of hybrid inviability. We found that incompatibilities between heterotypic cytoplasmic and nuclear loci commonly cause F2 inviability in hybridizations between wild-isolates of *Caenorhabditis nouraguensis*. We then mapped the nuclear incompatibility loci of two specific cytoplasmic nuclear incompatibilities, the NIC59 cytoplasmic-JU1825 nuclear and JU1825 cytoplasmic – NIC59 nuclear incompatibilities. Both incompatibilities are recessive, in that the nuclear incompatibility loci must be homozygous for their respective allele. We find that the JU1825 nuclear incompatibility locus maps to chromosome III and the NIC59 nuclear incompatibility locus maps to chromosome IV. We further narrowed down the JU1825 nuclear incompatibility locus to a repetitive 100 kb region. We attempted to generate NIC59 fosmid clones that span the candidate region to see which could rescue hybrid inviability. However, technical difficulties while generating the NIC59 fosmid library and potential selection against our fosmid clones of interest have halted our progress. We have currently isolated only one fosmid clone from the library that lies within the candidate region.

5.1.1 *Determining NIC59 loci required for viability*

To determine which NIC59 sequences in the candidate region are required for viability, we will isolate the remaining clones from the NIC59 fosmid library (sub-pooling first to get a sense of their actual frequency in the library) and then perform transgenic experiments to

determine whether some can rescue hybrid inviability. Generally, the rescue experiments will be performed in the following way. An injection mix composed of a linearized fosmid of interest, a linearized plasmid encoding a visible fluorescent marker and a linearized plasmid encoding neomycin resistance, will be injected into P0 NIC59 adult female germlines (anecdotally, I have found that injecting linearized plasmids increases the transformation frequency and therefore increases the chance of generating stable transgenic lines). Each injected female will be placed on their own NGM + OP50 plate and allowed to lay eggs for 1-2 days at 25°. Afterwards, neomycin will be added to the plate to select for F1 transformants. I would then look for stable lines in two to three generations (stable lines will starve their plates faster than unstable lines). Transgenic NIC59 females (as indicated by visible fluorescent marker) from a couple of transgenic lines will then be crossed to JU1825 males. Transgenic (N); N/J F1 females will then be crossed to JU1825 males. Individual viable adult F2 progeny will then be lysed and PCR genotyped to determine whether they carry the fosmid of interest and whether viable individuals carrying the fosmid can be homozygous for JU1825 alleles in the chromosome III candidate region. As a negative control, the same experiments will be done except using the original NIC59 strain that lacks the fosmid. Because the incompatibility depends on whether NIC59 or JU1825 is the mother in the first generation, it would be interesting to determine whether a fosmid's ability to rescue hybrid inviability is also dependent on which parent it was inherited from. For this reason, it would be interesting to also test fosmids that are inherited through the JU1825 male germline. This would be done by generating transgenic JU1825 lines and setting up the same crosses.

The fosmids that we have found in our current library do not completely span the candidate region. Furthermore, the fosmid we have isolated and end-sequenced appears to be a

chimera of two unlinked regions, further reducing our coverage of the candidate region. If none of the fosmid clones (or combinations of fosmid clones) rescue inviability, we might attempt to isolate additional fosmids that span the candidate region by generating a new library with higher coverage of the NIC59 genome. An alternative to generating a new fosmid library would be to clone and test whether individual NIC59 genes (or combinations of genes) can rescue hybrid inviability. This approach is impractical for several reasons: 1) We do not know how many NIC59 genes are required for viability and testing many combinations of genes will be labor intensive 2) The NIC59 genome is fragmented into many scaffolds in the candidate region. Therefore, some genes might be missing from our current assembly. Regardless of the approach taken, generating a more contiguous NIC59 assembly is essential to gain a better understanding of the differences between the candidate regions of NIC59 and JU2079 (inbred version of JU1825), and might give insight into the molecular mechanisms underlying the genetic incompatibility between these strains. PacBio sequencing in combination with our current Illumina reads will likely generate a more contiguous NIC59 assembly.

A complementary approach to determining which NIC59 regions are required for viability would be performing an association analysis of the candidate region between wild-isolates of *C. nouraguensis* that are either compatible or incompatible with the NIC59 cytoplasm. As mentioned in Chapter 2, currently all *C. nouraguensis* strains tested have a nuclear genome that is incompatible with the NIC59 cytoplasm (i.e. JU1837, JU1840, JU1854, NIC24 and NIC54). However, we currently do not know if these incompatibilities are genetically identical to the NIC59 cytoplasmic – JU1825 chromosome III incompatibility. To do so, we need to perform similar F1 female backcross experiments and test for selection against non-NIC59 alleles at the 100 kb candidate region on chromosome III by PCR genotyping. We can then compare the 100

kb candidate region between strains that show the same genetic incompatibility with the NIC59 cytoplasm with those that do not. This will require whole genome sequencing of additional *C. nouraguensis* wild-isolates.

Since NIC59 encodes something required for viability in the 100 kb candidate region that JU1825 does not, another approach to identifying that locus would be to make NIC59 deletions in that candidate region and see which result in embryonic lethality of (N); N/J F1 hybrids. Additionally, we can test the hypothesis that the transmembrane domain containing genes in the candidate region act as a cytoplasmic toxin by deleting them and seeing if F2 embryonic lethality is rescued.

5.1.2 *Determining the NIC59 cytoplasmic locus*

The (J); N/J F1 female x JU1825 male backcross experiments show there are two types of backcross lineages: those that begin to exhibit features of the NIC59 cytoplasmic – JU1825 nuclear incompatibility and those that do not. One hypothesis to explain this result is that JU1825 is naturally heteroplasmic and contains a NIC59-like mtDNA, which can increase in frequency in some backcross lineages and cause incompatibilities with JU1825 chromosome III. To test this hypothesis, we plan to deep sequence both types of backcross lineages and determine if there are associations between JU1825 mitochondrial variants and genetic incompatibility of JU1825 chromosome III.

If no associations are found, then it is unlikely that the mitochondrial genome harbors the NIC59 cytoplasmic incompatibility locus. We have also eliminated the possibility that maternally inherited endosymbiotic bacteria encode the cytoplasmic incompatibility locus. One possibility is that the maternally inherited NIC59 component of the incompatibility is not a

cytoplasmically inherited genome, but rather a nuclear NIC59 locus that is expressed in (N); N/J F1 female germlines but silenced in (J); N/J F1 female germlines. Thus, this nuclear NIC59 locus is expressed when inherited from females and silenced when inherited from males. This mechanism is also consistent with the (J); N/J F1 female x JU1825 male backcrossing data. Specifically, while (J); N/J F1 females largely do not express the nuclear NIC59 locus, her F2 progeny have inherited it through the female germline, and therefore can express it and cause incompatibility with JU1825 chromosome III. To test this hypothesis, we plan to perform RNAseq on bulk populations of (J); N/J and (N); N/J F1 females and look for genes that are significantly upregulated in (N); N/J F1 females.

5.2 RARE ASEXUAL REPRODUCTION IN *C. NOURAGUENSIS* INTRASPECIES CROSSES AND *C. NOURAGUENSIS* X *C. BECEI* HYBRIDIZATIONS

My second project focuses on understanding the evolutionary transition from sexual to asexual reproduction. We found that crossing two sexual species (*C. nouraguensis* females and *C. becei males*) results in rare gynogenetic reproduction. We find that gynogenetic reproduction results from diploid maternal inheritance (i.e. two random homologous chromatids from each maternal bivalent) and paternal genome loss. Additionally, we find that diploid maternal inheritance and sexual reproduction can occur in *C. nouraguensis* intraspecies crosses. We hypothesize that both interspecies and intraspecies diploid maternal inheritance can be used as stepping stones in the transition between sexual and asexual reproduction. Therefore, *C. nouraguensis* can be a useful model to study the characteristics of a nascent asexual lineage.

5.2.1 *Genetic variation for gynogenetic reproduction*

We have shown that the females of different wild-isolates of *C. nouraguensis* reproduce gynogenetically at significantly different rates when crossed to *C. becei* males. For example, crossing NIC54 *C. nouraguensis* females to QG711 *C. becei* males results in ~1/300 viable gynogenetic and hybrid offspring, whereas crossing JU1825 *C. nouraguensis* females to QG711 *C. becei* males results in ~1/1,000 viable gynogenetic and hybrid offspring. These results demonstrate that genetic differences between wild-isolates of *C. nouraguensis* influence the rate of interspecies gynogenesis. Mapping these genetic differences could give insight into the cellular mechanisms of rare gynogenetic reproduction in a primarily sexual species. Furthermore, population genetic studies could give insight into the evolutionary processes that have shaped the frequency of these loci within *C. nouraguensis*.

5.2.2 *Determining if rare gynogenesis occurs in other Caenorhabditis species*

We show that *C. nouraguensis* intraspecies crosses can undergo cryptic diploid maternal inheritance, an essential step towards the evolution of asexuality. Because diploid maternal inheritance occurs at such a low frequency and is probably not coupled with paternal genome loss, it would have gone undetected if not for the interspecies hybridization. Using UV irradiation to damage paternal DNA, we can now easily test how wide-spread cryptic diploid maternal inheritance occurs in other *Caenorhabditis* species. We will first test *C. elegans* because of the genetic and transgenic tools that can be used to study the genetic requirements and cell biology of rare asexuality. Most *Caenorhabditis* species appear to be obligately sexual, therefore finding many cases of cryptic diploid maternal inheritance could make the *Caenorhabditis* genus an ideal model to study the evolution of a nascent asexual lineage from a sexual ancestor.

5.2.3 *Disrupted female meiosis in dead F1 embryos*

We find that *C. becei* sperm disrupt *C. nouraguensis* female meiosis, resulting in the inheritance of at least two homologous chromatids in almost all dead F1 embryos. We hypothesize that *C. becei* sperm drastically increases the frequency of diploid maternal inheritance in *C. nouraguensis* oocytes, however we currently do not know the exact composition of the maternal genome in dead F1 embryos. To determine if dead F1 embryos undergo the same modified meiosis as viable F1, we plan to monitor female meiosis in real time using fluorescently tagged histone markers. If female meiosis is modified in the same way, then we predict that homologous chromosomes will segregate at anaphase I, but one set will fail to be segregated into a polar body. Then sister-chromatids would segregate at anaphase II, generating a single polar body. If female meiotic segregation in *C. nouraguensis* oocytes fertilized by *C. becei* sperm do not match these expectations, then perhaps the aberrant meiosis seen in dead embryos is unrelated to the diploid maternal inheritance seen in viable F1.

BIBLIOGRAPHY

Aalto, E. A., H.-P. Koelewijn, and O. Savolainen, 2013 Cytoplasmic Male Sterility Contributes to Hybrid Incompatibility Between Subspecies of Arabidopsis lyrata. G3 Genes|Genomes|Genetics 3: 1727–1740.

Aldrich, J. C., A. Leibholz, M. S. Cheema, J. Ausió, and P. M. Ferree, 2017 A "selfish" B chromosome induces genome elimination by disrupting the histone code in the jewel wasp Nasonia vitripennis. Sci. Rep. 7: 42551.

Andolfatto, P., D. Davison, D. Erezyilmaz, T. T. Hu, J. Mast *et al.*, 2011 Multiplexed shotgun genotyping for rapid and efficient genetic mapping. Genome Res. 21: 610–7.

Arnqvist, G., D. K. Dowling, P. Eady, L. Gay, T. Tregenza *et al.*, 2010 Genetic architecture of metabolic rate: Environment specific epistasis between mitochondrial and nuclear genes in an insect. Evolution (N. Y). 64: 3354–3363.

Baird, S. E., 2002 Haldane's Rule by Sexual Transformation in Caenorhabditis. Genetics 161: 1349–1353.

Baird, S. E., and R. Stonesifer, 2012 Reproductive isolation in Caenorhabditis briggsae: Dysgenic interactions between maternal- and zygotic-effect loci result in a delayed development phenotype. Worm 1: 189–95.

Baker, B. S., 1975 PATERNAL LOSS (PAL): A MEIOTIC MUTANT IN DROSOPHILA MELANOGASTER CAUSING LOSS OF PATERNAL CHROMOSOMES. Genetics 80: 267–96.

Baldo, L., J. C. D. Hotopp, K. A. Jolley, S. R. Bordenstein, S. A. Biber *et al.*, 2006 Multilocus sequence typing system for the endosymbiont Wolbachia pipientis. Appl. Environ. Microbiol. 72: 7098–7110.

Barbash, D. A., P. Awadalla, and A. M. Tarone, 2004 Functional divergence caused by ancient positive selection of a Drosophila hybrid incompatibility locus. PLoS Biol. 2: 839–848.

Beukeboom, L. W., and R. C. Vrijenhoek, 1998 Evolutionary genetics and ecology of sperm-dependent parthenogenesis. J. Evol. Biol. 11: 755–782.

Bhalla, N., and A. F. Dernburg, 2005 A conserved checkpoint monitors meiotic chromosome synapsis in Caenorhabditis elegans. Science 310: 1683–6.

Bi, Y., X. Ren, C. Yan, J. Shao, D. Xie *et al.*, 2015 A Genome-Wide Hybrid Incompatibility Landscape between Caenorhabditis briggsae and C. nigoni. PLoS Genet. 11: 1–26.

Bikard, D., D. Patel, C. Le Metté, V. Giorgi, C. Camilleri *et al.*, 2009 Divergent evolution of duplicate genes leads to genetic incompatibilities within A. thaliana. Science 323: 623–626.

Bordenstein, S. R., F. P. O'Hara, and J. H. Werren, 2001 Wolbachia-induced incompatibility precedes other hybrid incompatibilities in Nasonia. Nature 409: 707–710.

Bourtzis, K., A. Nirgianaki, G. Markakis, and C. Savakis, 1996 Wolbachia infection and cytoplasmic incompatibility in Drosophila species. Genetics 144: 1063–1073.

Brenner, S., 1974 The Genetics of Caenorhabditis elegans. Genetics 77: 71–94.

Brideau, N. J., H. a Flores, J. Wang, S. Maheshwari, X. Wang *et al.*, 2006 Two Dobzhansky-Muller genes interact to cause hybrid lethality in Drosophila. Science 314: 1292–1295.

Broman, K. W., H. Wu, S. Sen, and G. A. Churchill, 2003 R/qtl: QTL mapping in experimental crosses. Bioinformatics 19: 889–890.

Burke, N. W., and R. Bonduriansky, 2017 Sexual Conflict, Facultative Asexuality, and the True Paradox of Sex. Trends Ecol. Evol. 32: 646–652.

Burt, A., and R. Trivers, 2006 *Genes in conflict : the biology of selfish genetic elements*. Belknap Press of Harvard University Press.

Burton, R. S., and F. S. Barreto, 2012 A disproportionate role for mtDNA in Dobzhansky-Muller incompatibilities? Mol. Ecol. 21: 4942–4957.

Chae, E., K. Bomblies, S. T. Kim, D. Karelina, M. Zaidem *et al.*, 2014 Species-wide genetic incompatibility analysis identifies immune genes as hot spots of deleterious epistasis. Cell 159: 1341–1351.

Chang, C.-C., J. Rodriguez, and J. Ross, 2015 Mitochondrial-Nuclear Epistasis Impacts Fitness and Mitochondrial Physiology of Interpopulation Caenorhabditis briggsae Hybrids. G3 (Bethesda). 6: 209–19.

Chang, C.-C., C.-T. Ting, C.-H. Chang, S. Fang, and H.-Y. Chang, 2014 The Persistence of Facultative Parthenogenesis in Drosophila albomicans. PLoS One 9: e113275.

Chase, C. D., 2007 Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. Trends Genet. 23: 81–90.

Chou, J.-Y., Y.-S. Hung, K.-H. Lin, H.-Y. Lee, and J.-Y. Leu, 2010 Multiple molecular mechanisms cause reproductive isolation between three yeast species. PLoS Biol. 8: e1000432.

Clark, K. A., D. K. Howe, K. Gafner, D. Kusuma, S. Ping *et al.*, 2012 Selfish little circles: Transmission bias and evolution of large deletion-bearing mitochondrial DNA in caenorhabditis briggsae nematodes. PLoS One 7: 1–8.

Collins, T. J., 2007 ImageJ for microscopy. Biotechniques 43: S25–S30.

Corbett-Detig, R. B., J. Zhou, A. G. Clark, D. L. Hartl, and J. F. Ayroles, 2013 Genetic incompatibilities are widespread within species. Nature 504: 135–7.

Coyne, J. A., and H. A. Orr, 2004 *Speciation*. Sinauer Associates, Sunderland, MA.

Coyne, J. a, and H. a Orr, 1998 The evolutionary genetics of speciation. Philos. Trans. R. Soc.

Lond. B. Biol. Sci. 353: 287–305.

Cutter, A. D., 2012 The polymorphic prelude to Bateson-Dobzhansky-Muller incompatibilities. Trends Ecol. Evol. 27: 209–218.

Cutter, A. D., S. E. Baird, and D. Charlesworth, 2006 High nucleotide polymorphism and rapid decay of linkage disequilibrium in wild populations of Caenorhabditis remanei. Genetics 174: 901–913.

Cutter, A. D., A. Dey, and R. L. Murray, 2009 Evolution of the Caenorhabditis elegans genome. Mol. Biol. Evol. 26: 1199–1234.

Cutter, A. D., R. Jovelin, and A. Dey, 2013 Molecular hyperdiversity and evolution in very large populations. Mol. Ecol. 22: 2074–2095.

D'Souza, T. G., and N. K. Michiels, 2010 The costs and benefits of occasional sex: Theoretical predictions and a case study. J. Hered. 101: 34–41.

Dacks, J., and A. J. Roger, 1999 The First Sexual Lineage and the Relevance of Facultative Sex. J. Mol. Evol. 48: 779–83.

Darby, A. C., A. C. Gill, S. D. Armstrong, C. S. Hartley, D. Xia *et al.*, 2015 Integrated transcriptomic and proteomic analysis of Wolbachia to doxycycline-induced stress. Mol. Cell. Proteomics 14: 1038–53.

Davies, A. B., E. Hatton, N. Altemose, J. G. Hussin, F. Pratto *et al.*, 2016 Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice. Nature 530: 171–176.

Davis, M. P. A., S. van Dongen, C. Abreu-Goodger, N. Bartonicek, and A. J. Enright, 2013 Kraken: A set of tools for quality control and analysis of high-throughput sequence data. Methods 63: 41–49.

Delph, L. F., and J. P. Demuth, 2016 Haldane's rule: Genetic bases and their empirical support.

J. Hered. 107: 383–391.

Dey, A., C. K. W. Chan, C. G. Thomas, and A. D. Cutter, 2013 Molecular hyperdiversity defines populations of the nematode Caenorhabditis brenneri. Proc. Natl. Acad. Sci. U. S. A. 110: 11056–60.

Dey, A., Y. Jeon, G.-X. Wang, and A. D. Cutter, 2012 Global population genetic structure of Caenorhabditis remanei reveals incipient speciation. Genetics 191: 1257–1269.

Dey, A., Q. Jin, Y. C. Chen, and A. D. Cutter, 2014 Gonad morphogenesis defects drive hybrid male sterility in asymmetric hybrid breakdown of Caenorhabditis nematodes. Evol. Dev. 16: 362–372.

Dobzhansky, T., 1933 On the sterility of the interracial hybrids in Drosophila pseudoobscura. Proc. Natl. Acad. Natl. Acad. Sci. 19: 397–403.

Dörner, M., M. Altmann, S. Pääbo, and M. Mörl, 2001 Evidence for import of a lysyl-tRNA into marsupial mitochondria. Mol. Biol. Cell 12: 2688–2698.

Edgar, R. C., 2004 MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32: 1792–1797.

Eisman, R., and T. C. Kaufman, 2007 Cytological Investigation of the Mechanism of Parthenogenesis in Drosophila mercatorum. Fly 1: 317–329.

Ellison, C. K., and R. S. Burton, 2008 Interpopulation hybrid breakdown maps to the mitochondrial genome. Evolution (N. Y). 62: 631–638.

Ellison, C. K., O. Niehuis, and J. Gadau, 2008 Hybrid breakdown and mitochondrial dysfunction in hybrids of Nasonia parasitoid wasps. J. Evol. Biol. 21: 1844–1851.

Engelstädter, J., 2008 Constraints on the evolution of asexual reproduction. BioEssays 30: 1138–1150.

English, A. C., S. Richards, Y. Han, M. Wang, V. Vee *et al.*, 2012 Mind the Gap: Upgrading Genomes with Pacific Biosciences RS Long-Read Sequencing Technology (Z. Liu, Ed.). PLoS One 7: e47768.

Félix, M. A., C. Braendle, and A. D. Cutter, 2014 A streamlined system for species diagnosis in caenorhabditis (Nematoda: Rhabditidae) with name designations for 15 distinct biological species. PLoS One 9:.

Félix, M.-A., R. Jovelin, C. Ferrari, S. Han, Y. R. Cho *et al.*, 2013 Species richness, distribution and genetic diversity of Caenorhabditis nematodes in a remote tropical rainforest. BMC Evol. Biol. 13: 10.

Felsenstein, J., 1976 THE EVOLUTIONARY ADVANTAGE OF RECOMBINATION. Genetics 78: 737–756.

Ferree, P. M., and D. A. Barbash, 2009 Species-specific heterochromatin prevents mitotic chromosome segregation to cause hybrid lethality in Drosophila. PLoS Biol. 7:.

Fields, A. T., K. A. Feldheim, G. R. Poulakis, and D. D. Chapman, 2015 Facultative parthenogenesis in a critically endangered wild vertebrate. Curr. Biol. 25: R446–R447.

Fierst, J. L., J. H. Willis, C. G. Thomas, W. Wang, R. M. Reynolds *et al.*, 2015 Reproductive Mode and the Evolution of Genome Size and Structure in Caenorhabditis Nematodes. PLoS Genet. 11: e1005323.

Fischer, G., S. a James, I. N. Roberts, S. G. Oliver, and E. J. Louis, 2000 Chromosomal evolution in Saccharomyces. Nature 405: 451–454.

Fopp-Bayat, D., K. Ocalewicz, M. Kucinski, M. Jankun, and B. Laczynska, 2017 Disturbances in the ploidy level in the gynogenetic sterlet Acipenser ruthenus. J. Appl. Genet. 58: 373–380.

Fujii, S., C. S. Bond, and I. D. Small, 2011 Selection patterns on restorer-like genes reveal a

conflict between nuclear and mitochondrial genomes throughout angiosperm evolution. Proc. Natl. Acad. Sci. U. S. A. 108: 1723–1728.

Fuyama, Y., 1986a Genetics of Parthenogenesis in DROSOPHILA MELANOGASTER. I. The Modes of Diploidization in the Gynogenesis Induced by a Male-Sterile Mutant, ms(3)K81. Genetics 112: 237–248.

Fuyama, Y., 1986b Genetics of Parthenogenesis in DROSOPHILA MELANOGASTER. II. Characterization of a Gynogenetically Reproducing Strain. Genetics 114: 495–509.

Gaborieau, L., G. G. Brown, and H. Mireau, 2016 The Propensity of Pentatricopeptide Repeat Genes to Evolve into Restorers of Cytoplasmic Male Sterility. Front. Plant Sci. 7: 1816.

Gemmell, N. J., V. J. Metcalf, and F. W. Allendorf, 2004 Mother's curse: The effect of mtDNA on individual fitness and population viability. Trends Ecol. Evol. 19: 238–244.

Gibeaux, R., R. Acker, M. Kitaoka, G. Georgiou, I. van Kruijsbergen *et al.*, 2018 Paternal chromosome loss and metabolic crisis contribute to hybrid inviability in Xenopus. Nature 553: 337–341.

Gibson, A. K., L. F. Delph, and C. M. Lively, 2017 The two-fold cost of sex: Experimental evidence from a natural system. Evol. Lett. 1: 6–15.

Gitschlag, B. L., C. S. Kirby, D. C. Samuels, R. D. Gangula, S. A. Mallal *et al.*, 2016 Homeostatic Responses Regulate Selfish Mitochondrial Genome Dynamics in C. elegans. Cell Metab. 24: 91–103.

Goda, T., A. Abu-Daya, S. Carruthers, M. D. Clark, D. L. Stemple *et al.*, 2006 Genetic screens for mutations affecting development of Xenopus tropicalis. PLoS Genet. 2: e91.

Goddard, K. A., O. Megwinoff, L. L. Wessner, and F. Giaimo, 1998 Confirmation of gynogenesis in Phoxinus eos-neogaeus (Pisces: Cyprinidae). J. Hered. 89: 151–157.

Gorchs Rovira, A., and A. G. Smith, 2019 PPR proteins - orchestrators of organelle RNA metabolism. Physiol. Plant.

Green, R. F., and D. L. G. Noakes, 1995 Is a little bit of sex as good as a lot? J. Theor. Biol. 174: 87–96.

Greenfield, P., K. Duesing, A. Papanicolaou, and D. C. Bauer, 2014 Blue: correcting sequencing errors using consensus and context. Bioinformatics 30: 2723–2732.

Groot, T. V. M., E. Bruins, and J. A. J. Breeuwer, 2003 Molecular genetic evidence for parthenogenesis in the Burmese python, Python molurus bivittatus. Heredity (Edinb). 90: 130–135.

Grosmaire, M., C. Launay, M. Siegwald, T. Brugière, L. Estrada-Virrueta *et al.*, 2019 Males as somatic investment in a parthenogenetic nematode. Science 363: 1210–1213.

Guindon, S., J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk *et al.*, 2010 New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst. Biol. 59: 307–321.

Gustafsson, C. M., M. Falkenberg, and N.-G. Larsson, 2016 Maintenance and Expression of Mammalian Mitochondrial DNA. Annu. Rev. Biochem. 85: annurev-biochem-060815-014402.

Hicks, K. A., D. K. Howe, A. Leung, D. R. Denver, and S. Estes, 2012 In Vivo Quantification Reveals Extensive Natural Variation in Mitochondrial Form and Function in Caenorhabditis briggsae. PLoS One 7:.

Hiruta, C., C. Nishida, and S. Tochinai, 2010 Abortive meiosis in the oogenesis of parthenogenetic Daphnia pulex. Chromosom. Res. 18: 833–840.

Holt, C., and M. Yandell, 2011 MAKER2: an annotation pipeline and genome-database

management tool for second-generation genome projects. BMC Bioinformatics 12: 491.

Huang, W., C. Yu, J. Hu, L. Wang, Z. Dan *et al.*, 2015 Pentatricopeptide-repeat family protein RF6 functions with hexokinase 6 to rescue rice cytoplasmic male sterility. Proc. Natl. Acad. Sci. U. S. A. 112: 14984–9.

Huber, W., V. J. Carey, R. Gentleman, S. Anders, M. Carlson *et al.*, 2015 Orchestrating high-throughput genomic analysis with Bioconductor. Nat. Methods 12: 115–121.

Hunt, M., T. Kikuchi, M. Sanders, C. Newbold, M. Berriman *et al.*, 2013 REAPR: a universal tool for genome assembly evaluation. Genome Biol. 14: R47.

Hunter, N., S. R. Chambers, E. J. Louis, and R. H. Borts, 1996 The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. TL - 15. EMBO J. 15 VN-r: 1726–1733.

Hunter, S. S., R. T. Lyon, B. A. J. Sarver, K. Hardwick, L. J. Forney *et al.*, 2015 Assembly by Reduced Complexity (ARC): a hybrid approach for targeted assembly of homologous sequences. bioRxiv 014662. doi: 10.1101/014662.

Itono, M., N. Okabayashi, K. Morishima, T. Fujimoto, H. Yoshikawa *et al.*, 2007 Cytological Mechanisms of Gynogenesis and Sperm Incorporation in Unreduced Diploid Eggs of the Clonal Loach, Misgurnus anguillicaudatus (Teleostei: Cobitidae). J. Exp. Zool. 307A: 35–50.

Jaquiéry, J., S. Stoeckel, C. Larose, P. Nouhaud, C. Rispe *et al.*, 2014 Genetic Control of Contagious Asexuality in the Pea Aphid. PLoS Genet. 10: e1004838.

Kajitani, R., K. Toshimoto, H. Noguchi, A. Toyoda, Y. Ogura *et al.*, 2013 Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. Genome Res. 24: 1384–1395.

Kassir, Y., D. Granot, and G. Simchen, 1988 IME1, a Positive Regulator Gene of Meiosis in S. cerevisiae. Cell 52: 853–862.

Kent, W. J., R. Baertsch, A. Hinrichs, W. Miller, and D. Haussler, 2003 Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes. Proc. Natl. Acad. Sci. U. S. A. 100: 11484–9.

Kiontke, K. C., M.-A. Félix, M. Ailion, M. V. Rockman, C. Braendle *et al.*, 2011 A phylogeny and molecular barcodes for Caenorhabditis, with numerous new species from rotting fruits. BMC Evol. Biol. 11: 339.

van der Kooi, C. J., and T. Schwander, 2015 Parthenogenesis: Birth of a New Lineage or Reproductive Accident? Curr. Biol. 25: R654–R676.

Kozlowska, J. L., A. R. Ahmad, E. Jahesh, and A. D. Cutter, 2012 Genetic variation for postzygotic reproductive isolation between caenorhabditis briggsae and caenorhabditis sp. 9. Evolution (N. Y). 66: 1180–1195.

Kurtz, S., A. Phillippy, A. L. Delcher, M. Smoot, M. Shumway *et al.*, 2004 Versatile and open software for comparing large genomes. Genome Biol. 5: R12.

Laetsch, D. R., and M. L. Blaxter, 2017 BlobTools: Interrogation of genome assemblies. F1000Research 6: 1287.

Lamatsch, D. K., M. Schmid, and M. Schartl, 2002 A somatic mosaic of the gynogenetic Amazon molly. J. Fish Biol. 60: 1417–1422.

Lamelza, P., and M. Ailion, 2017 Cytoplasmic–Nuclear Incompatibility Between Wild Isolates of Caenorhabditis nouraguensis. G3 7: 823–834.

Lamelza, P., J. M. Young, L. M. Noble, A. Isakharov, M. Palanisamy *et al.*, 2019 Cryptic asexual reproduction in Caenorhabditis nematodes revealed by interspecies hybridization.

bioRxiv 588152.

Laser, K. D., and N. R. Lersten, 1972 Anatomy and cytology of microsporogenesis in cytoplasmic male sterile angiosperms. Bot. Rev. 38: 425–454.

Lee, H. Y., J. Y. Chou, L. Cheong, N. H. Chang, S. Y. Yang *et al.*, 2008 Incompatibility of Nuclear and Mitochondrial Genomes Causes Hybrid Sterility between Two Yeast Species. Cell 135: 1065–1073.

Leggett, R. M., B. J. Clavijo, L. Clissold, M. D. Clark, and M. Caccamo, 2014 NextClip: an analysis and read preparation tool for Nextera Long Mate Pair libraries. Bioinformatics 30: 566–568.

Lemire, B., 2005 Mitochondrial genetics. WormBook.

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

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Liénard, M. A., L. O. Araripe, and D. L. Hartl, 2016 Neighboring genes for DNA-binding proteins rescue male sterility in *Drosophila* hybrids. Proc. Natl. Acad. Sci. 201608337.

Liu, Y., J. Schröder, and B. Schmidt, 2013 Musket: a multistage k-mer spectrum-based error corrector for Illumina sequence data. Bioinformatics 29: 308–315.

Long, Y., L. Zhao, B. Niu, J. Su, H. Wu *et al.*, 2008 Hybrid male sterility in rice controlled by interaction between divergent alleles of two adjacent genes. Proc. Natl. Acad. Sci. 105: 18871–18876.

Luo, D., H. Xu, Z. Liu, J. Guo, H. Li *et al.*, 2013 A detrimental mitochondrial-nuclear interaction causes cytoplasmic male sterility in rice. Nat. Genet. 45: 573–577.

Ma, H., N. Marti Gutierrez, R. Morey, C. Van Dyken, E. Kang *et al.*, 2015 Incompatibility between Nuclear and Mitochondrial Genomes Contributes to an Interspecies Reproductive Barrier. Cell Metab. 1–12.

Mayr E., 1942 *Systematics and the Origin of Species*, Columbia University Press, New York.

Madl, J. E., and R. K. Herman, 1979 POLYPLOIDS AND SEX DETERMINATION IN CAENORHABDITIS ELEGANS. Genetics 93: 393–402.

Maheshwari, S., and D. A. Barbash, 2011 The Genetics of Hybrid Incompatibilities. Annu. Rev. Genet. 45: 331–355.

Markow, T. A., 2013 Parents Without Partners: Drosophila as a Model for Understanding the Mechanisms and Evolution of Parthenogenesis. G3 3: 757–762.

Masly, J. P., C. D. Jones, M. A. F. Noor, J. Locke, and H. A. Orr, 2006 Gene transposition as a cause of hybrid sterility in Drosophila. Science 313: 1448–1450.

Maynard Smith, J., 1971 What use is sex? J. Theor. Biol. 30: 319–335.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis *et al.*, 2010 The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20: 1297–1303.

McNally, K. P., M. T. Panzica, T. Kim, D. B. Cortes, and F. J. McNally, 2016 A novel chromosome segregation mechanism during female meiosis. Mol. Biol. Cell 27: 2576–2589.

McOrist, S., 2000 Obligate intracellular bacteria and antibiotic resistance. Trends Microbiol. 8: 483–486.

Meiklejohn, C. D., M. A. Holmbeck, M. A. Siddiq, D. N. Abt, D. M. Rand *et al.*, 2013 An Incompatibility between a Mitochondrial tRNA and Its Nuclear-Encoded tRNA Synthetase

Compromises Development and Fitness in Drosophila. PLoS Genet. 9:.

Meneely, P. M., A. F. Farago, and T. M. Kauffman, 2002 Crossover distribution and high interference for both the X chromosome and an autosome during oogenesis and spermatogenesis in Caenorhabditis elegans. Genetics 162: 1169–1177.

Mihola, O., Z. Trachtulec, C. Vlcek, J. C. Schimenti, and J. Forejt, 2009 A Mouse Speciation Gene Encodes a Meiotic Histone H3 Methyltransferase. Science (80-.). 323: 373–5.

Mirzaghaderi, G., and E. Hörandl, 2016 The evolution of meiotic sex and its alternatives. Proc. R. Soc. B Biol. Sci. 283: 20161221.

Murdy, W. H., and H. L. Carson, 1959 Parthenogenesis in Drosophila mangabeirai Malog. Am. Nat. 93: 355–363.

Nelson-Rees, W. A., M. A. Hoy, and R. T. Roush, 1980 Heterochromatinization, chromatin elimination and haploidization in the parahaploid mite Metaseiulus occidentalis (Nesbitt) (Acarina: Phytoseiidae). Chromosoma 77: 263–276.

Okonechnikov, K., O. Golosova, and M. Fursov, 2012 Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28: 1166–1167.

Oliveira, D. C. S. G., R. Raychoudhury, D. V. Lavrov, and J. H. Werren, 2008 Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp Nasonia (Hymenoptera: Pteromalidae). Mol. Biol. Evol. 25: 2167–2180.

Oliver, P. L., L. Goodstadt, J. J. Bayes, Z. Birtle, K. C. Roach *et al.*, 2009 Accelerated evolution of the Prdm9 speciation gene across diverse metazoan taxa. PLoS Genet. 5:.

Olshen, A. B., E. S. Venkatraman, R. Lucito, and M. Wigler, 2004 Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 5: 557–572.

Orr, H. A., 1995 The population genetics of speciation: The evolution of hybrid

incompatibilities. Genetics 139: 1805–1813.

Osada, N., and H. Akashi, 2012 Mitochondrial-nuclear interactions and accelerated compensatory evolution: Evidence from the primate cytochrome c oxidase complex. Mol. Biol. Evol. 29: 337–346.

Palopoli, M. F., M. V. Rockman, A. TinMaung, C. Ramsay, S. Curwen *et al.*, 2008 Molecular basis of the copulatory plug polymorphism in Caenorhabditis elegans. Nature 454: 1019–1022.

Phadnis, N., E. P. Baker, J. C. Cooper, K. A. Frizzell, E. Hsieh *et al.*, 2015 An essential cell cycle regulation gene causes hybrid inviability in Drosophila. Science 350: 1552–5.

Phadnis, N., and H. A. Orr, 2009 A Single Gene Causes Sterility and Segregation Distortion in Drosophila hybrids. Science (80-.). 323: 376–379.

Pires-daSilva, A., 2007 Evolution of the control of sexual identity in nematodes. Semin. Cell Dev. Biol. 18: 362–370.

Presgraves, D. C., 2010 The molecular evolutionary basis of species formation. Nat. Rev. Genet. 11: 175–180.

Presgraves, D. C., L. Balagopalan, S. M. Abmayr, and H. A. Orr, 2003 Adaptive evolution drives divergence of a hybrid inviability gene between two species of Drosophila. Nature 423: 715–719.

Ragavapuram, V., E. E. Hill, and S. E. Baird, 2016 Suppression of F1 Male-Speci fi c Lethality in Caenorhabditis Hybrids by cbr-him-8. 6: 623–629.

Rand, D. M., R. A. Haney, and A. J. Fry, 2004 Cytonuclear coevolution: The genomics of cooperation. Trends Ecol. Evol. 19: 645–653.

Robinson, J. T., H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander *et al.*, 2011

Integrative genomics viewer. Nat. Biotechnol. 29: 24–26.

Rockman, M. V., and L. Kruglyak, 2009 Recombinational landscape and population genomics of Caenorhabditis elegans. PLoS Genet. 5: e1000419.

Roelens, B., M. Schvarzstein, and A. M. Villeneuve, 2015 Manipulation of karyotype in Caenorhabditis elegans reveals multiple inputs driving pairwise chromosome synapsis during meiosis. Genetics 201: 1363–1379.

Ross, J. A., D. K. Howe, A. Coleman-hulbert, D. R. Denver, and S. Estes, 2016 Paternal Mitochondrial Transmission in Intra-Species Caenorhabditis briggsae Hybrids. 33: 3158–3160.

Ross, J. A., D. C. Koboldt, J. E. Staisch, H. M. Chamberlin, B. P. Gupta *et al.*, 2011 Caenorhabditis briggsae recombinant inbred line genotypes reveal inter-strain incompatibility and the evolution of recombination. PLoS Genet. 7: e1002174.

Sadler, P. L., and D. C. Shakes, 2000 Anucleate Caenorhabditis elegans sperm can crawl, fertilize oocytes and direct anterior-posterior polarization of the 1-cell embryo. Development 127: 355–366.

Sambatti, J. B. M., D. Ortiz-Barrientos, E. J. Baack, and L. H. Rieseberg, 2008 Ecological selection maintains cytonuclear incompatibilities in hybridizing sunflowers. Ecol. Lett. 11: 1082–1091.

Sanei, M., R. Pickering, K. Kumke, S. Nasuda, and A. Houben, 2011 Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. Proc. Natl. Acad. Sci. 108: E498–E505.

Schwander, T., S. Vuilleumier, J. Dubman, and B. J. Crespi, 2010 Positive feedback in the transition from sexual reproduction to parthenogenesis. Proc. R. Soc. B Biol. Sci. 277:

1435–1442.

Seidel, H. S., M. Ailion, J. Li, A. van Oudenaarden, M. V. Rockman *et al.*, 2011 A novel sperm-delivered toxin causes late-stage embryo lethality and transmission ratio distortion in C. elegans. PLoS Biol. 9:.

Seidel, H. S., M. V Rockman, and L. Kruglyak, 2008 Widespread genetic incompatibility in C. elegans maintained by balancing selection. Science 319: 589–594.

Severson, A. F., L. Ling, V. van Zuylen, and B. J. Meyer, 2009 The axial element protein HTP-3 promotes cohesin loading and meiotic axis assembly in C. elegans to implement the meiotic program of chromosome segregation. Genes Dev. 23: 1763–1778.

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

Simpson, J. T., and R. Durbin, 2012 Efficient de novo assembly of large genomes using compressed data structures. Genome Res. 22: 549–56.

Skaar, J. R., J. K. Pagan, and M. Pagano, 2013 Mechanisms and function of substrate recruitment by F-box proteins. Nat. Rev. Mol. Cell Biol. 14: 369–381.

Stalker, H. D., 1954 Parthenogenesis in Drosophila. Genetics 39: 4–34.

Steinbrenner, A. D., S. Goritschnig, K. V Krasileva, K. J. Schreiber, and B. J. Staskawicz, 2012 Effector recognition and activation of the Arabidopsis thaliana NLR innate immune receptors. Cold Spring Harb. Symp. Quant. Biol. 77: 249–57.

Stelkens, R. B., C. Schmid, and O. Seehausen, 2015 Hybrid breakdown in cichlid fish. PLoS One 10: 1–11.

Stelzer, C.-P., and J. Lehtonen, 2016 Diapause and maintenance of facultative sexual

reproductive strategies. Philos. Trans. R. Soc. B Biol. Sci. 371: 20150536.

Stevens, L., M.-A. Félix, T. Beltran, C. Braendle, C. Caurcel *et al.*, 2018 Comparative genomics of ten new Caenorhabditis species. bioRxiv 398446. doi: 10.1101/398446.

Sunnucks, P., and D. F. Hales, 1996 Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus Sitobion (Hemiptera: Aphididae). Mol. Biol. Evol. 13: 510–524.

Tao, Y., D. L. Hartl, and C. C. Laurie, 2001 Sex-ratio segregation distortion associated with reproductive isolation in Drosophila. Proc. Natl. Acad. Sci. U. S. A. 98: 13183–8.

Taylor, J., and D. Butler, 2017 R Package ASMap: Efficient Genetic Linkage Map Construction and Diagnosis. J. Stat. Softw. 79: 1–29.

Templeton, A. R., H. L. Carson, and C. F. Sing, 1976 The population genetics of parthenogenetic strains of Drosophila mercatorum. II. The capacity for parthenogenesis in a natural, bisexual population. Genetics 82: 527–542.

Ting, C.-T., S.-C. Tsaur, M.-L. Wu, and C.-I. Wu, 1998 A Rapidly Evolving Homeobox at the Site of a Hybrid Sterility Gene. Science (80-.). 282: 1501–1504.

Turelli, M., and L. C. Moyle, 2007 Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. Genetics 176: 1059–1088.

Wallace, D. C., and D. Chalkia, 2013 Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. Cold Spring Harb. Perspect. Biol. 5:.

Watts, P. C., K. R. Buley, S. Sanderson, W. Boardman, C. Ciofi *et al.*, 2006 Parthenogenesis in Komodo dragons. Nature 444: 1021–1022.

Werren, J. H., L. Baldo, and M. E. Clark, 2008 Wolbachia: master manipulators of invertebrate biology. Nat. Rev Microbiol 6: 741–751.

Woodruff, G. C., O. Eke, S. E. Baird, M. A. Félix, and E. S. Haag, 2010 Insights into species divergence and the evolution of hermaphroditism from fertile interspecies hybrids of Caenorhabditis nematodes. Genetics 186: 997–1012.

Wu, T. D., and S. Nacu, 2010 Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics 26: 873–881.

Yamaki, T., G. K. Yasuda, and B. T. Wakimoto, 2016 The deadbeat paternal effect of uncapped sperm telomeres on cell cycle progression and chromosome behavior in Drosophila melanogaster. Genetics 203: 799–816.

Yasuda, G. K., G. Schubiger, and B. T. Wakimoto, 1995 Genetic Characterization of ms(3) K81, a paternal effect gene of Drosophila melanogaster. Genetics 140: 219–229.

Zanders, S. E., M. T. Eickbush, J. S. Yu, J. W. Kang, K. R. Fowler *et al.*, 2014 Genome rearrangements and pervasive meiotic drive cause hybrid infertility in fission yeast. Elife 3: e02630.

VITA

Piero Lamelza received a B.S. in Molecular, Cellular and Developmental Biology at the University of California Santa Cruz (UCSC) in Summer 2010. While attending UCSC he worked as an undergraduate research assistant in Susan Strome's laboratory for a year at UCSC and also as a summer intern research assistant in Priscilla K. Cooper's lab at Lawrence Berkeley National Laboratory. After graduating from UCSC, he worked as a research technician in Needhi Bhalla's laboratory at UCSC for two years where he studied female meiosis in *C. elegans*.