ABSTRACT

Title of Dissertation: GENETIC AND MOLECULAR ANALYSIS OF GERMLINE SEX DETERMINATION IN *CAENORHABDITIS BRIGGSAE*, A MODEL FOR THE CONVERGENT EVOLUTION OF HERMAPHRODITISM

Alana V. Doty, Doctor of Philosophy, 2010

Directed By: Dr. Eric S. Haag, Associate Professor
Department of Biology

Though sex determination and differentiation are critical biological processes, genetic mechanisms that specify sex have undergone profound and rapid evolutionary change across taxa. We may be able to infer processes that generate sex determination diversity by examining closely related species. Within the nematode genus *Caenorhabditis*, two species, *C. elegans* and *C. briggsae*, are androdioecious, producing self-fertile hermaphrodites and males; other *Caenorhabditis* species generate males/females. Interestingly, phylogenies reveal that *C. elegans* and *C. briggsae* independently acquired hermaphroditism, a relatively rare adaptation among animals. In this work, I describe differences in germline sex determination between *C. elegans* and *C. briggsae* that may help reveal the molecular basis of their convergent evolution of hermaphroditism. I first describe mutations in the pleiotropic, STAR family RNA-binding protein Cbr-GLD-1 that affect germline sex in *C. briggsae*. I find that *C. briggsae gld-1* mutant hermaphrodites have a sex determination phenotype opposite to that of *C. elegans*: masculinized versus feminized germlines. I demonstrate that Cbr-GLD-1 coding-plus-regulatory sequences can rescue *Ce-gld-1* null animals, arguing that this change in sex determination is not due to changes in GLD-1 function or expression. I further show that *gld-1*'s role in regulating oogenesis is conserved across the Elegans group of *Caenorhabditis*, demonstrating that the

oogenesis function of *gld-1* is likely ancient, whereas its sperm-repressing role in *C. briggsae* has evolved recently. To identify mRNA targets of Cbr-GLD-1 that might be responsible for its sex determination function in *C. briggsae*, I use an *in vivo* genome-wide approach to isolate mRNAs associated with Cbr-GLD-1, including potential sex determination targets. I identify 800 putative mRNA targets and confirm specificity of this gene set via qRT-PCR and RNAi. Next, to reveal the roles of GLD-1 in evolutionary context, I create a phylogeny of STAR proteins across metazoans. Finally, I characterize a single feminizing allele recovered through forward screens in *C. briggsae* for germline sex determination mutants. This work thus begins to dissect the molecular and genetic basis of hermaphroditism in *C. briggsae* and contributes to a growing body of research on the evolution of germline sex determination differences between *C. elegans* and *C. briggsae*.

Genetic and molecular analysis of germline sex determination in *Caenorhabditis briggsae*, a model for the convergent evolution of hermaphroditism

By

Alana V. Doty

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2010

Advisory Committee:

Associate Professor Eric Haag, *Chair*

Associate Professor Alexandra Bely

Professor William Jeffery

Associate Professor Leslie Pick

Dr. Harold Smith

Dedication

For Grandma, Mot, and Pop

Acknowledgments

I would like to first thank my advisor, Eric Haag, for his support. Dr. Haag not only provided funding and scientific guidance for this work, but also filled a vital role as its unfailing cheerleader. His enthusiasm and kindness were of inestimable value while traversing what can be the long, winding road of graduate school. I have not only become a better scientist, but also a better person, while in his laboratory.

I would also like to thank members of the Haag lab I have had the pleasure of working with and for helpful discussions and advice: Danielle, Robin, Gavin, Qinwen, Cris, Joe, Mandy, Kevin, John S., Andy, Brian, Shanni, Dorothy, Onyi, and other rotation (e.g., Siqian, Yizhou) and visiting (e.g., Te-Wen, John W.) students. Indeed, labmates are more than just coworkers, and though perhaps my productivity suffered at the hands of our many conversations, I learned and laughed enough to make them worth it. In this vein, I would also like to thank my BEES and larger Biology fellow graduate students for their support, social events, and for making our program a positive place in which to spend time and develop as researchers. I hope I have helped contribute to our Program as well.

Lastly, I would like to thank my immediate and extended family for their continuous understanding and support. My mother and father are parents without parallel, and I cannot possibly imagine how my fiancé Jon has stuck around after putting up with so many of my work-related travails. My extended family, housemates, and siblings were always good for reminding me that even if one's experiments don't work, there are always sausage and pepper sandwiches to eat, movies to watch, and vacations to take - i.e., life is still sweet.

Table of Contents

List of Tables

CHAPTER 1

CHAPTER 3

List of Figures

INTRODUCTION

CHAPTER 1

numbers in parentheses below the X-axis present the 'number of germlines scored / number of animals', as for some animals both gonad arms were scored independently. The "Pro tumor" and "some sperm" categories are not mutually exclusive.

FIGURE 8. An amino acid alignment of Cbr-GLD-1 and Ce-GLD-1. *Cbr-gld-1* mutations are indicated with black boxes and font (*nm41, nm64*, and *nm68*), and the corresponding *Ce-gld-1* mutations are in orange font (as well as the *q93* mutation that contributes to *q93oz50*); the *Ce-gld-1* null allele *q485* is also shown. Lighter blue shading is for conservative amino acid substitutions, and dark blue designates identical amino acids; no shading indicates more radical amino acid substitutions and insertions/deletions.

FIGURE 9. Western blot analysis of wild-type *C. elegans* and *C. briggsae* hermaphrodites and XX animals homozygous for the *C. elegans gld-1* null allele *q485*, *C. briggsae* deletion allele *nm68*, *C. briggsae* nonsense mutations *nm41* and *nm64*, and the corresponding mutations in *C. elegans q268* and *q93oz50. C. elegans q485* and *q268* mutants possess a germline tumor ("Tum"), and the *q93oz50* mutant phenotype is sperm and cells in meiotic arrest. *C. briggsae* mutants are as described in text. Protein is probed with an antibody to the oocyte yolk receptor RME-2 and an antibody to the Major Sperm Protein, MSP, in order to assess the sexual fate of mutant germlines in the absence of true gametogenesis. The anti-MSP staining in *Cbr-gld-1* mutants likely only derives, however, from mature sperm and the few recognizable spermatocytes, not from grossly tumorous cells (Figures 10 and 11). Tubulin is a loading control; *C. briggsae gld-1* mutants reproducibly produce more tubulin than the same number of wild-type animals.

FIGURE 10. *Cbr-gld-1(nm68)* dissected gonads stain with an antibody to Major Sperm Protein. (A) Image from Lints and Hall (2009). DAPI-stained dissected gonad from a wild-type *C. elegans* hermaphrodite to illustrate *Caenorhabditis* germline chromosome morphology. Image indicates the progression of germ cells from the self-renewing germ cell source, the "mitotic zone," to fully differentiated gametes (oocytes and sperm); (B) dissected *nm68* gonad arm stained with Hoechst-33258 to identify chromosome morphology. Though the mitotic region and transition zone are normal, the worm overproduces sperm compared to wild-type animals and has spermatocytes even in adulthood. At the proximal end of the gonad is ectopic germline proliferation, and just distal to the tumor is a region of mixed pachytene-like cells and mitotic cells; (C) same gonad arm as in (B), stained with anti-MSP antibody, a marker for the sperm fate. anti-MSP stains primary spermatocytes and sperm, but not mitotic cells; (D) same gonad arm as in (C), stained with an antibody to the oocyte yolk receptor RME-2 as a marker for oocyte fate. anti-RME-2-stained image is overexposed to detect potential weak staining.

FIGURE 11. *Cbr-gld-1(nm41)* dissected gonads stain with an antibody to Major Sperm Protein. (A) Hoechst-stained *nm41* dissected gonad to reveal chromosome morphology. Only cells proximal to the transition zone are shown. This adult gonad still contains primary spermatocytes and overproduces sperm. Gonad also has a region of mixed pachytene and mitotic cells; (B) and (C) as in Figure 11. anti-MSP antibody stains mature sperm and primary spermatocytes. anti-RME-2-stained image is overexposed to detect potential weak staining.

FIGURE 12. *Cbr-gld-1*(nm64) dissected gonads display both MSP and weak RME-2 staining. (A) Hoechst-stained *nm64* dissected gonad arm showing a normal mitotic region but an abnormally large transition zone and cells with unusual DNA morphology. As in Figures 10 and 11, this gonad arm has an over-proliferation of sperm compared to wild-type animals and still contains spermatocytes; (B) and (C) as in Figures 10 and 11. (B) anti-MSP antibody stains only the mature sperm; (C) *nm64* gonad arms sometimes exhibit weak RME-2 staining in meiotic-like cells near the transition zone. anti-RME-2-stained image is overexposed to detect potential weak staining.

FIGURE 13. *nm64* dissected gonad stained with Hoechst (left) and anti-RME-2 antibody (right). RME-2 is detected in cells proximal to the transition zone with abnormal chromosome morphology. anti-RME-2-stained image is overexposed to detect potential weak staining.

FIGURE 14. *C. elegans gld-1*(q485) null dissected gonad arm stained with (B) Hoechst dye for chromosome morphology, (C) anti-MSP antibody to mark the sperm fate, and (D) anti-RME-2 antibody to mark the oocyte fate.

FIGURE 15. *Ce-gld-1*(q485 null) homozygotes are rescued by *C. briggsae gld-1*. Top panel depicts a *C. elegans gld-1* null XX animal, with ectopic germ cell proliferation and no sperm. Bottom panel depicts a *Ce-gld-1* null animal rescued to self-fertility by pAD-g6, containing the *Cbr-gld-1* locus. All obvious *q485* defects have been rescued, including oogenic meiotic progression, specification of hermaphrodite sperm, and formation of the germline rachis. (DTC is the distal tip cell of the somatic gonad.)

FIGURE 16. Western blot showing both GLD-1 and HA expression in the two *Ce-gld-1*(q485 null) rescued strains CP113 and CP114. Wild type *C. elegans* and *C. briggsae* hermaphrodites have abundant GLD-1 but no HA protein, and *Ce-gld-1* null homozygotes have neither GLD-1 nor HA. Blots were first probed with rabbit anti-GLD-1, then stripped and reprobed with anti-tubulin (loading control), then stripped again and probed with anti-HA. Each lane contains 40 animals.

FIGURE 17. Confirmation that self-fertile *unc-13* homozygous animals are *Ce-gld-1*(q485) homozygotes in CP113 and CP114. The *q*485 allele is an 82bp deletion at the *gld-1* locus. *Ce-gld-1* was amplified from *q485* heterozygotes and homozygotes and from Uncs of the strain BS3156 rescued to self-fertility in CP113 or CP114.

FIGURE 18. Western blot analysis of GLD-1 in females/hermaphrodites (karyotype XX) and males (karyotype XO) from 7 species of *Caenorhabditis*, including undescribed species Sp. 5 and Sp 9. anti-GLD-1 antibody is chicken 1026 anti-GLD-1, described in dissertation Chapter 2; tubulin is a "loading control." Crude protein was made from 40 adults per lane. Given above species names is a cladogram depicting the evolutionary relationships of taxa. Thick blue bars indicate a gain of hermaphroditism in *C. elegans* and *C. briggsae*.

FIGURE 19. DIC micrographs of XX-karyotype F1 progeny from *gld-1* RNAi injected mothers of (A) Sp. 9–*C. briggsae* hybrid; (B) *C. remanei*; (C) *C. brenneri* ; (D) *C. japonica*; (E) *C. briggsae*. All germlines except *C. briggsae* develop a germline tumor and no sperm, whereas *C. briggsae* affected germlines contain only sperm. White triangles indicate the proximal end of the gonad, near the vulva.

CHAPTER 2

FIGURE 1. Schematic of microarray design for two different experiments, anti-GLD-1 IP vs. anti-IgY IP mRNA expression comparison and anti-GLD-1 IP vs. total input mRNA expression comparison, on two Agilent 4 x 44K arrays. Dye swaps and biological replicates are incorporated.

FIGURE 2. Images from Qinwen Liu (unpublished results). Chicken anti-GLD-1 antibody recognizes native GLD-1. Fixed, dissected gonad of a *C. briggsae* wild-type hermaphrodite shown (A) with differential interference microscopy (DIC), (B) Hoechst-stained to observe DNA morphology, and (C) stained with chicken anti-GLD-1. The Cbr-GLD-1 pattern here is the same as observed by Nayak et al. (2005).

FIGURE 3. Work flow for identifying Cbr-GLD-1 mRNA targets by RIP-Chip.

FIGURE 4. 1% TBE-agarose gel electrophoresis of total RNA obtained from worm lysis/pre-clearing steps ("total RNA") and after immunoprecipitation ("S/N RNA") in the RIP-Chip protocol. Each letter (F-K) designates a biological replicate. The bright RNA bands are ribosomal RNA, and their integrity indicates a general lack of RNA degradation.

FIGURE 5. Western blot of immunoprecipitation material in two biological replicates. Samples were taken from pre-cleared input material before immunoprecipitation ("total lysate"), from post-incubation antibody-bound beads, and from immunoprecipitation supernatants after incubation. The anti-GLD-1 antibody used in this assay recognizes different epitopes from that used for the immunoprecipitations. "Mock" refers to total IgY-bound agarose beads.

FIGURE 6. Individual array boxplots of (A) *normexp* background corrected, eCADS normalized intensities for the anti-GLD-1 IP vs. anti-IgY IP microarray comparison, and (B) no background-corrected, eCADS normalized intensities for the anti-GLD-1 IP vs. total mRNA comparison.

FIGURE 7. Venn diagram (Oliveros 2007) of two microarray comparisons, anti-GLD-1 IP mRNA vs. mock anti-IgY IP mRNA and anti-GLD-1 IP mRNA vs. total input mRNA, each analyzed with two differential gene expression programs, SAM and EDGE. Values in each oval are the number of probes enriched in anti-GLD-1 immunoprecipitations with FDRs of at most <2% (except see * in diagram). 2,966 probes were found in common to all four data sets, representing 802 *C. briggsae* protein-coding genes.

FIGURE 8. Quantitative RT-PCR assessment of Cbr-GLD-1 immunoprecipitated candidate sex determination targets. Enrichment is calculated by dividing the inferred amount of starting material in anti-GLD-1 IPs by that in mock anti-IgY IPs for each gene, averaged over at least 3 biological replicates. Additionally, though we could amplify *Cbr-fog-3* from total mRNA material, we could not detect it in either mRNA recovered from anti-GLD-1 IPs nor from mock anti-IgY IPs for four replicates.

FIGURE 9. Alignment of representative metazoan STAR domains. The three conserved regions are shown in sequence order: QUA1 (involved in homodimerization, except in the SF1 subfamily that remains as monomers; their sequences are more divergent here) (Zorn and Kreig 1997, Chen et al. 1997, Liu et al. 2001, Beuck et al. 2010), the KH RNA-binding domain, and the QUA2 domain (which provides at least an extended RNA-binding surface) (Liu et al. 2001, Ryder et al. 2004, Maguire et al. 2005). In addition, STAR proteins form protein contacts with different binding partners (for instance, Taylor et al. 1994, Clifford et al. 2000, Selenko et al. 2003, Najib 2005, Robard et al. 2006). Alignment was performed with ClustalX 2.0.12 according to default parameters; residue are colored according to ClustalX defaults to highlight chemically similar amino acids.

FIGURE 10. Bayesian phylogenetic tree of the STAR protein family in representative metazoans. The tree is rooted at the SF1 clade of ancient splicing factors. Values along the tree backbone and at intermediate nodes are posterior probabilities; for visual clarity, terminal node posterior probabilities are not given, but values are similarly high to those shown. The subfamilies of well-studied proteins are highlighted in color: GLD-1 and the related ASD-2 in dark green and light green, respectively; How/Who in purple; Quaking in orange; SF1 in red; and SAM68 and the related SLM-1 and -2 in dark blue and light blue, respectively. Proteins are named with their genus/species abbreviation (see key) and then with either their NCBI Protein sequence database name (like "Quaking"), their NCBI GI identification number, or for *Caenorhabditis* sequences, their WormBase protein ID.

CHAPTER 3

FIGURE 1. Forward mutagenesis scheme. Young adult wild-type *C. briggsae* AF16 hermaphrodites are exposed to 50mM EMS for 4 hours at room temperature. F1 self-progeny are picked to fresh plates and allowed to self. Recessive Fog alleles are recovered in the F2.

FIGURE 2. *nm38* mutants are Fog. (A) presumed *nm38* XX young adult animal in early gametogenesis. The first germ cells to develop are oocytes, and the inset highlights an empty spermathecae and an oocyte as

the most proximal gamete; (B) the same animal as in (A) recovered overnight to a fresh plate. She laid no self-progeny, still possesses no sperm, and displays the stacking oocyte Fog phenotype; (C) young adult wild type-looking sibling of (A) clearly undergoing spermatogenesis; (D) *nm38* XO male with a perfect male soma but both sperm and oocytes in the germline.

FIGURE 3. Western blot of *C. briggsae* AF16 wild-type and *nm38* mutant animals. Protein in the top panel was exposed to an antibody to the oocyte yolk-receptor RME-2. Protein in the bottom panel was exposed to an antibody to the Major Sperm Protein. Middle panel is tubulin loading control. * Only half as many AF16 males are loaded compared to the XX lanes.

FIGURE 4. *nm38* intersexual somas. (A), (B), (C) are F1 from heterozygote sib-crosses; (D) F1 from AF16 XX x *nm38* XO cross.

FIGURE 5. Presumed *nm38 ed23*ts double mutant reared at the restrictive temperature with clear oocytes in the germline instead of only sperm.

FIGURE 6. Fingerprint contig map of *C. briggsae* chromosome 2 from WormBase (build CB3). Colored arrows mark the locations of particular AF16-VT847 SNP mapping assays; identically colored arrows indicate assays predicted to lie on the same fpc. The numbers above each arrow are the number of map units of assay from *nm38*.

INTRODUCTION:

Sex determination mechanisms are widely diverged across taxa

Sex determination and differentiation are critical for many aspects of biology, affecting the developmental programs of organisms, their morphology and behavior, and the population genetic dynamics of their species. However, the mechanisms that specify sex determination undergo profound and rapid evolutionary change (for instance, Organ et al. 2009, Mank and Avise 2009, Wallis et al. 2008, Barske and Capel 2008, Bull 1983). Indeed, marked variation exists in both primary sex determination signals and in the downstream genetic pathways that interpret these signals. Broadly, sex determination systems are divided into two mechanistic categories: environmental sex determination, in which environmental features such as temperature or local sex ratio dictate sexual differentiation, and genetic sex determination, in which the sex is set by a sex chromosome or an autosomal gene(s). Though little is known about the molecular mechanisms of environmental sex determination (Janzen and Phillips 2006, Valenzuela 2008), many different genetic mechanisms of sex determination have been uncovered. For instance, sex chromosome systems in which either the female (ZW/ZZ) or the male (XX/XY) is heterogametic, are common, as are systems set by the ratio of X chromosomes to sets of autosomes. There are also systems in which heterozygosity at a single locus is required for sexual development, as well as those involving multiple genes with additive effects (e.g., Bull 1983, chapter 2 & 3).

Past these primary sex determination systems, we find a great variety of downstream genetic pathways that interpret these signals. For example, although sex is initially set in both *D. melanogaster* and *C. elegans* by the X:A ratio, the *Drosophila* pathway consists of a cell-autonomous cascade of regulated mRNA splicing, while *Caenorhabditis* uses a Hedgehog-like

signaling pathway (reviewed in Cline and Meyer 1996). In contrast, sex determination in most mammals depends upon a male-specific transcription factor (Sry) that is invoked early in development for testes formation (reviewed in Sinclair et al. 2002). In birds, sex was similarly assumed to be controlled through the testis as in mammals, but now avian somatic cells are though to possess a cell autonomous, inherent sexual identity of their own (Zhao et al. 2010).

What genetic and evolutionary processes cause sex determination mechanisms to change so dramatically? To understand the evolution of sex determination, we may be able to identify evolutionary steps by examining differences between closely related species. Looking over a small evolutionary window may allow us to capture intermediate genetic changes not visible between more widely diverged taxa.

Nematode study system

The nematode genus *Caenorhabditis* is an excellent system in which to study the evolution of sex determination. This is true for two reasons, the first of which is that the sex determination pathway of the model species *C. elegans*, among the first animal developmental pathways to be characterized, is now known in great genetic and biochemical detail. The thorough understanding of *C. elegans'* sex determination facilitates characterization of congeners' sex determination systems (Haag 2005, Pires-daSilva 2007). Secondly, *Caenorhabditis* is unique among the genera of biological model system in possessing species with different mating systems. We can study both the convergent evolution of the rarest of all animal mating systems, androdioecy, in which a species produces both hermaphrodites and males (Weeks et al. 2006, Jarne and Auld 2006); and also the repeated evolution of androdioecy from gonochoristic (male and female producing) ancestors within this genus.

There are 10 described species of *Caenorhabditis*, and most of them are gonochoristic. However, two are androdioecious: the model species *C. elegans* and satellite model *C. briggsae*. Somatic and germline anatomy are shared between *C. elegans* and *C. briggsae* hermaphrodites and males, though the two species diverged 20-30 million years ago (Figure 1: Cutter 2008).

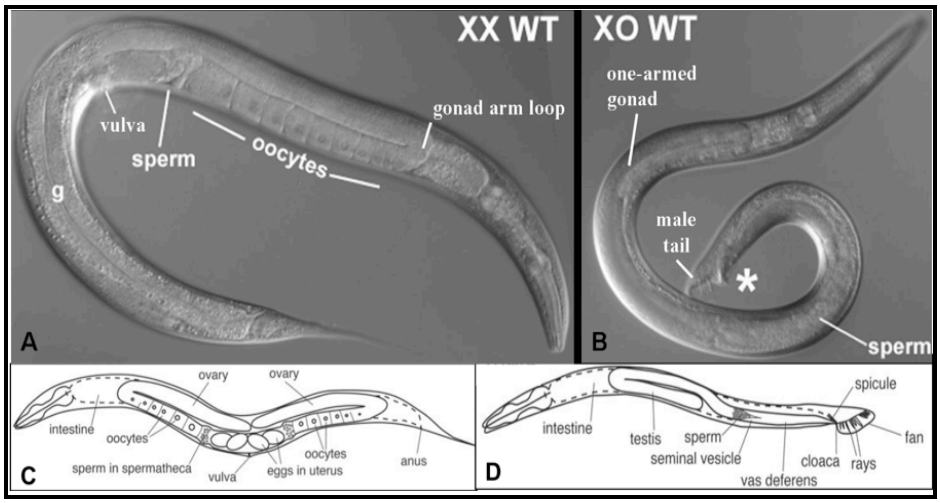


FIGURE 1. (A) and (B) DIC micrographs of a wild-type adult C. briggsae XX hermaphrodite and XO male. Some sexually dimorphic characters are indicated. Image provided by Eric Haag. (C) and (D) Cartoons illustrating some adult sexually dimorphic features of *Caenorhabditis elegans* hermaphrodites and males, respectively. Image taken from Zarkower 2006.

Hermaphrodites of both species (XX karyotype) possess a fully female soma, but make sperm transiently at the time of their last larval molt; then, just a few hours later in development, they switch to producing oocytes from the same population of germ cells for the remainder of their lives. Both *C. elegans* and *C. briggsae* hermaphrodites are capable of either selfing or of using male sperm for fertilization. Males of both species (XO karyotype) are produced naturally by sex chromosome nondisjunction at a low frequency. Males possess a specialized somatic anatomy for mating and make sperm continuously as adults.

Surprisingly, phylogenetic analysis indicates that *C. elegans* and *C. briggsae* independently evolved androdioecy from different male/female ancestral species (Figure 2: Kiontke et al. 2004). A growing body of genetic evidence revealing differences in germline sex

determination between *C. elegans* and *C. briggsae* is consistent with this finding (Nayak et al. 2005, Hill et al. 2006, Guo et al. 2009).

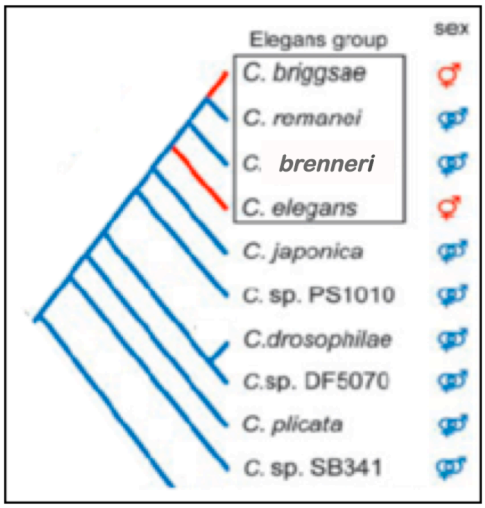


FIGURE 2. Modified from Kiontke et al., 2004. A cladogram of cultured *Caenorhabditis* species. The mating system of each taxa is indicated, blue for gonochoristic and red for androdioecious species.

Androdioecy within *Caenorhabditis* is thus an example of convergent evolution, the process by which independent, and potentially unique, changes in a developmental pathway can lead to the same overt phenotypic evolutionary change.

Present study

The goal of this work is to uncover molecular and genetic changes in germline sex determination between *C. elegans* and *C. briggsae* in order to help understand their convergent evolution of an important and relatively rare adaptation among animals: hermaphroditism.

In the first chapter, I describe the recovery and characterization of mutations in the conserved, pleiotropic STAR family RNA-binding protein Cbr-GLD-1 that affect germline sex determination in *C. briggsae*. *gld-1* has orthologs throughout *Caenorhabditis*, and some aspects of GLD-1 function are conserved between *C. elegans* and *C briggsae*. Surprisingly, however, GLD-

1 has an opposite major effect on sex determination in the two species: loss of *Ce-gld-1* feminizes the hermaphrodite germline, whereas loss of *Cbr-gld-1* masculinizes the hermaphrodite germline. We find that the Cbr-GLD-1 can rescue the *Ce-gld-1* null phenotype, arguing that *gld-1*'s different sex determination effects are due to differences in the network of protein or mRNA factors with which GLD-1 interacts, not in GLD-1 itself.

In the second chapter, I continue to characterize *C. briggsae gld-1* and seek to identify its messenger RNA targets, including those involved in sex determination. I do this by immunoprecipitating GLD-1 from *C. briggsae* adult hermaphrodites and use microarrays to detect transcripts enriched in anti-GLD-1 immunoprecipitations. I confirm the specificity of recovered putative Cbr-GLD-1 targets with qRT-PCR for positive and negative control genes and by assaying for RNAi phenotypes of 100 of these targets. I fail to recover sex determination phenotypes from these RNAi injections, however, and also do not detect binding of Cbr-GLD-1 to candidate sex determination genes from either microarray analysis or by assaying directly with qRT-PCR. To further understand Cbr-GLD-1's functions and mechanism of action, we placed GLD-1 in its larger evolutionary context and constructed a phylogeny of STAR-domain proteins across metazoans. We find that GLD-1 is part of a clade of nematode-specific proteins, and is most closely related to well-studied STAR family members *Drosophila* How/Who and vertebrate Quaking.

In the third chapter, I characterize a single feminizing allele recovered through forward screens in *C. briggsae* hermaphrodites for germline sex determination mutants. This mutation, *nm38*, feminizes both the germlines of hermaphrodites and males, and can cause somatic feminization of XO animals in certain crosses. We find that *nm38* is not allelic to other known *C. briggsae* feminizing genes, and that it can suppress spermatogenesis in two other sex determination mutant alleles. We map *nm38* to the right end of *C. briggsae* chromosome 2.

CHAPTER 1:

Mutations in *C. briggsae gld-1* affect germline sex determination and germline development

ABSTRACT

Though sex determination and differentiation are critical for many aspects of biology, the mechanisms that specify and implement sexual fates are among the least-conserved of developmental processes. Given the great diversity of sex determination mechanisms that exist, we may be able to infer changes that lead to such diversity by examining closely related species. Within the genus *Caenorhabditis*, *C. elegans* and *C. briggsae* are androdioecious, producing self-fertile hermaphrodite and males; other *Caenorhabditis* species generate males and females. Surprisingly, phylogenies indicate that *C. elegans* and *C. briggsae* independently acquired the ability to produce hermaphrodites from different male/female ancestral species. In this work, we investigate the independent evolution of hermaphroditism in *C. elegans* and *C. briggsae* and describe the first genetic mutants that masculinize the germline of *C. briggsae* hermaphrodites. These alleles all affect *C. briggsae* GLD-1, an RNA binding protein of the STAR-domain family. We find, in concordance with an earlier RNAi-based study (Nayak et al. 2005), that *C. briggsae gld-1* mutant hermaphrodites have a sex determination phenotype opposite that of *C. elegans*. We demonstrate that Cbr-GLD-1 can fully rescue *Ce-gld-1* null animals, arguing that this change in sex determination function is not due to changes in GLD-1 function or expression. We further show that *gld-1*'s role in regulating oogenesis is conserved across the Elegans group of *Caenorhabditis*, demonstrating that the oogenesis function of *gld-1* is likely ancient, whereas its sperm-repressing role in *C. briggsae* has evolved recently, perhaps in concert with the evolution of selfing in this lineage.

INTRODUCTION

Sex determination and differentiation are critical for many aspects of biology, affecting organisms' developmental programs, morphology and behavior, and population genetic dynamics. Most animals produce two sexes, and phylogenies suggest that the production of sperm and oocytes is an ancient metazoan character. However, though critical for development and reproduction, the mechanisms that specify and implement sexual fates are among the least-conserved of developmental processes (for instance, Bull 1983, Graves 2008, Barske and Capel 2008, Meiklejohn and Tao 2010).

Given the great diversity of sex determination mechanisms that exist across taxa, we can infer molecular changes that lead to such diversity by examining closely related species, among which sex determination changes are still interpretable. Within the genus *Caenorhabditis*, two described species, *C. elegans* and *C. briggsae*, are androdioecious, producing self-fertile hermaphrodite and male sexes; other *Caenorhabditis* species generate males and females. Surprisingly, *Caenorhabditis* phylogenies indicate, and genetic evidence supports, that *C. elegans* and *C. briggsae* independently acquired the ability to produce hermaphrodites from different male/female ancestral species (Kiontke et al. 2004, Cho et al. 2004, Nayak et al. 2005, Hill et al. 2006, Guo et al. 2009).

Somatic and germline anatomy are shared among the hermaphrodites and males of *C. elegans* and *C. briggsae*, though the two species may have diverged 20-30 million years ago (Cutter 2008). Hermaphrodites of *C. elegans* and *C. briggsae* are essentially females that are able to make sperm transiently near the time of their last larval molt; they can use either cross sperm from true males or self sperm for fertilization. *C. elegans* and *C. briggsae* males possess a specialized somatic anatomy for mating and make sperm continuously as adults.

Importantly, among the first animal developmental pathways to be described, the sex determination pathway of *C. elegans* is known in great detail (reviewed in Kimble and Crittenden 2007, Ellis 2008). This understanding facilitates characterization of other nematode sex determination pathways and investigations into the evolution of sex determination (Haag 2005, Pires-daSilva 2007).

In this work, we investigate the independent evolution of hermaphroditism in *Caenorhabditis* and describe the first genetic mutants that masculinize the germline of *C. briggsae* hermaphrodites. These alleles all affect *C. briggsae* GLD-1, an RNA binding protein of the STAR family (for **s**ignal **t**ransduction and **a**ctivation of **R**NA metabolism). STAR proteins have been implicated in a diverse set of cellular processes, and family members are found across metazoans (e.g., Volk et al. 2008, Galarneau and Richard 2005, Lukong and Richard 2003, Lee and Schedl 2001, Arning et al. 1996, dissertation Chapter 2). In *C. elegans*, GLD-1 is a pleiotropic, germline regulator known to promote hermaphrodite spermatogenesis by translationally repressing the female-promoting *tra-2* mRNA (Francis at al. 1995a and 1995b, Jones et al. 1996, Jan et al. 1999) .

In this work, we find, in concordance with an earlier RNAi-based study (Nayak et al. 2005), that *C. briggsae gld-1* mutants have a mutant sex determination phenotype opposite that of *C. elegans*, causing only sperm to be made in XX hermaphrodites. We demonstrate that *Cbr-gld-1* can rescue *Ce-gld-1* null animals, revealing that this change in sex determination function between *C. elegans* and *C. briggsae gld-1* is not likely due to changes in the GLD-1 protein or its expression. We further show that *gld-1*'s role in regulating oogenesis is conserved across the Elegans group of *Caenorhabditis.* Finally, we characterize the aberrant meiosis and tumors in *Cbr-gld-1* mutant germlines, and we posit that they might result from a novel role of *gld-1* in hermaphrodite spermatogenesis.

METHODS

Forward genetic screens

We performed forward mutant screens using 50mM EMS on synchronous *C. briggsae* young hermaphrodites (just after their L4-adult molt) nutating for 4 hours at room temperature. This P_0 generation was extensively washed in M9 and plated on standard NGM plates at 20°C. We singled F1 hermaphrodite L4s two to a 6cm plate and let them lay F2 progeny at either 20°C or 25°C (for warm temperature-sensitive allele screens). Putative Mog mutants were identified by first screening F2 plates with a dissecting microscope for adult sterility. About 10% of total F2 plates contained sterile animals without oocytes, and from these, multiple animals per plate were then examined with Nomarski optics for overproduction of sperm. Mog strains were maintained by sib-selection and were outcrossed at least six times to AF16 males before characterization.

Deletion mutation screens

Reverse genetic screens for targeted *Cbr-gld-1* deletion mutants in AF16 were performed as described in Hill et al. 2006. Screening of 2,000,000 haploid genomes, we recovered a single mutation, *nm68*, which removes 923 base pairs from the *gld-1* coding region. This allele was outcrossed to wild-type AF16 *C. briggsae* six times (twice through the mother) prior to characterization.

DNA sequencing

We used Big Dye v3.1 Cycle Sequencing chemistry (Applied Biosystems) for DNA sequencing according to the manufacturer's instructions. Sequencing was performed on ABI 3100 or 3750 machines according to standard protocols, and trace files were examined with the

"4 Peaks" program (Mekentosj software) and analyzed or further manipulated with Vector NTI software (Invitrogen).

SNP genotyping

We developed SNP-based linkage assays that take advantage of polymorphisms between the *C. briggsae* mutagenesis strain AF16 and mapping strain HK104. Sequencing within HK104, we identified a 90 bp deletion in the largest intron of AF16 *atx-2* relative to HK104 and also an RFLP in the DNA immediately downstream of *gld-1* in AF16 compared to HK104. We mated *nm41*/+ and *nm64*/+ hermaphrodites to HK104 wild-type males and then singled and selfed virgin hermaphrodite F1 progeny of successful crosses. F2 sterile progeny were genotyped, and F1 selfing mothers were also genotyped to ensure that she was the product of outcrossing.

Small-scale RNA preps

50 worms of specific genotypes were picked into nuclease-free water in 1.5ml microfuge tubes. About 5 volumes of Tri Reagent Solution (Ambion) was added, and tubes were frozen at -80°C. At a later date, the tubes were thawed at room temperature and spun down to pellet worms. An RNase-free plastic pestle (USA-Scientific, #1415-5390) was inserted slowly so as to not disturb the pellet, and worms were ground for about 20 seconds. Worms were again spun to the bottom of the tube and grinding was repeated. A third round of grinding was often necessary for male somas. (The grinding progress can be checked by looking for broken worm bodies with a dissecting microscope.) Next, Trizol was added to 1ml, and then 5ul polyacryl carrier (Molecular Research Center) added. Tubes were mixed well and incubated at room temperature for 5 minutes. Finally, 200ul chloroform was added to each tube, and the RNA

extracted and precipitated with 500ul isopropanol and washed according to standard procedures.

RT-PCR

Small-scale prep RNA (usually 2.5ul/reaction, corresponding to 2.5 worms) and gene-specific primers were used in Promega's Access RT-PCR kit according to manufacturer's instructions. Products were visualized on agarose/TBE gels.

RNAi

We used WormBase (WS210) to obtain DNA sequences for predicted genes. For the *Caenorhabditis*-wide *gld-1* RNAi experiment, we designed primers to amplify the same conserved ~750bp region of *gld-1* in each species (and in paralogs, if applicable), corresponding to nucleotides 753-1626 of *C. elegans gld-1*. Templates for production of double stranded RNA were produced by PCR from genomic DNA preps for each species with T7 promoter-tailed primers. PCR product templates were used in the Megascript T7 kit (Ambion) to produce double-stranded RNA according to manufacturer's protocol. RNA was purified according to the ammonium acetate cleanup procedure and resuspended in TE. We injected double stranded RNA at a concentration of ~3ug/ul into the gut of middle aged XX adults who had been grown at 20°C. Injected mothers were moved to fresh plates every 12-24 hours at 20°C and their progeny examined by DIC when adults. For injections into gonochoristic species, we placed injected mothers together with conspecific males, even if plugged prior to injection, to facilitate laying a large brood.

Western blotting

Western blotting was performed according to standard procedures. 40 animals of a particular genotype or sex were picked into a small amount of PBS in the cap of a microfuge tube and spun down, washed once in PBS, then resuspended with an equal volume of standard 2X protein sample buffer with 5% BME. Tubes were heated to 95°C for 5 minutes and then frozen at -80°C until use. Nitrocellulose membranes were blocked in 5% non-fat dried milk in PBS with 0.1% Tween-20, and antibodies were diluted in the same. Anti-RME-2 antibody was a gift of Dr. Barth Grant (Rutgers University; Grant and Hirsh 1998), and was used at 1:2000. Anti-MSP was a gift of Dr. David Greenstein (University of Minnesota; Kosinski et al. 2005) and was used at 1:5000. Anti-tubulin antibody (Sigma T9026) was used at 1:2000. We also used 3 anti-GLD-1 antibodies: the first, a gift of Tim Schedl (Washington University of St. Louis) was made against the C-terminal 82% of *C. elegans* GLD-1; it also detects *C. briggsae* GLD-1. The second antibody is a gift of Judith Kimble (University of Wisconsin, Madison; antibody produced by the former laboratory of Elizabeth Goodwin); it was made against *C. elegans* GLD-1 and also cross-reacts in *C. briggsae*. The final antibody was produced in our own laboratory against a 20 amino acid peptide near the N-terminus of *C. briggsae* GLD-1; it is described in dissertation Chapter 2.

Gonad dissection and antibody staining

We largely followed protocols for dissecting gonads, fixation, and antibody/DAPI (or Hoechst dye) staining from the laboratory of Dr. Tim Schedl, using the methanol/formaldehyde fix for 10-15 minutes. We modified these protocols by manipulating worms in low retention microfuge tubes (instead of glass tubes) and then washing/ blocking the dissected gonads in 4x15minute washes in PBST + 0.1% BSA post fixation. Primary and secondary antibodies were diluted in PBS + 0.1% BSA, and we included two extra wash steps post secondary antibody

incubation. Primary antibodies were those described above, and we used anti-MSP at 1:4000 and anti-RME-2 and anti-GLD-1 at 1:000–1:2000. Fluorescently-conjugated secondary antibodies were Alexa 488 and Alexa 555 (Molecular Probes-Invitrogen), used at a 1:1000-1:2000 dilution. All dissected gonads were blocked, incubated in primary and secondary antibodies, and washed simultaneously in the same conditions.

Multi-site Gateway **Cloning** ***(Invitrogen)***

Based upon the WormBase gene prediction, we used PCR and the Gateway BP reaction to clone 923bp upstream of the AF16 *Cbr-gld-1* start codon into the Multi-Site Gateway vector pDONR P4-P1R and 1354bp downstream of the *Cbr-gld-1* stop codon into the Multi-Site Gateway vector pDONR P2R-P3 according to manufacturer's instructions.

We used a C-terminal epitope tag for Cbr-GLD-1 as in Lee et al. (2001). We PCR-amplified the entire *Cbr-gld-1* coding region, including introns and start codon, using forward primer attB1F for cloning into pDONR 221 and a reverse primer containing the HA tag (adjusted for preferred *C. elegans* codons) and the last 21bp of *Cbr-gld-1*. To double-tag Cbr-GLD-1, we amplified a very small primer-based product using forward primer with both the HA and FLAG sequences and reverse primer containing the FLAG and attB2R sequences for cloning into pDONR 221. Then we “sewed” these two gel-purified PCR products together using their common HA tags in a PCR reaction using equimolar ratios of 50:1 *gld-1*-HA : HA-FLAG as template, Turbo Pfu polymerase from Stratagene (Agilent Technologies), and an annealing time of 1 minute in 50 cycles. These conditions were used to give the rare template (the HA-FLAG PCR product) very little opportunity to anneal to itself and to instead encourage formation of *gld-1*-HA/HA-FLAG annealed DNA fragments; adding the PCR primers attB1F and attB2R from the beginning of the “sewing” PCR reaction also likely encourages full length product formation.

Finally, we gel purified the correct-length PCR product from this reaction, verified by it sequencing, and performed the pDONR221 BP reaction according to Invitrogen instructions.

Colonies from the three successful BP reactions (*Cbr-gld-1* upstream, coding sequence+tags, and downstream region) were restriction digested and sequenced to identify correct donor plasmids. These plasmids were combined together for the Multi-Site Gateway LR reaction according to manufacturer's instructions to produce a single plasmid, pAD-g1.

To create our final bombardment vector, we excised the total *Cbr-gld-1* fragment from pAD-g1 by restriction digestion and subcloned it into pCR50, a gift of Christopher Richie (laboratory of Andrew Golden, NIDDK-NIH). pCR50 contains an *unc-119* rescuing fragment (from the *Cbr-unc-119* gene) as well as a *Ce-myo-2*::GFP transgene (originally from pOK100.03, Peter Okkema) that contains a derived promoter sequence to drive strong GFP expression in the worm pharynx. The resulting plasmid is pAD-g6 (Figure 1).

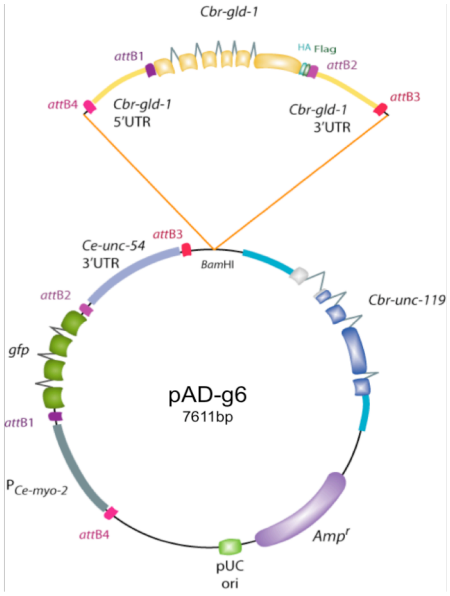


FIGURE 1. (Figure created by Cristel Thomas.) A cartoon of pAD-g6, used to rescue the null allele *Ce-gld-1*(q485). See text for construction details.

Bombardment

We bombarded pAD-g6 into *C. elegans* generally following the protocols of the labs of Geraldine Seydoux (Johns Hopkins Medical School) and Barth Grant (Rutgers University) on a bombardment apparatus kindly supplied by Iqbal Hamza (University of Maryland). We grew the standard *C. elegans* bombardment strain *Ce-unc-119*(ed3) at 20°C on NGM plates seeded with *E. coli* strain HB101, a thicker-growing relative of OP-50. We then harvested 1ml of gravid animals, bombarded them with 10ug of plasmid, and recovered animals to 80 10cm NGM plates seeded with HB101 at 25°C. After 10-12 days, we screened worms for both wild-type moving and GFP-pharynx expressing animals; using two selection markers increases the chance that observed transgenic worms are stable integrants. We identified 19 plates with >3 such animals and singled 4 animals/plate to maintain these "lines" and to check for continued transgene expression.

RESULTS

nm41, nm64, and nm68 are alleles of C. briggsae gld-1

We conducted forward genetic screens in the canonical, sequenced *C. briggsae* strain AF16 for recessive alleles that masculinize the hermaphrodite germline (so-called Mog alleles for **m**asculinization **o**f **g**ermline) according to the scheme in Figure 2. Screening 3000 haploid genomes, we recovered two alleles, *nm41* and *nm64*, that have a range of mutant sterile germline phenotypes including germline masculinization. As described below, these are alleles of *Cbr-gld-1*. We then sought a deletion allele in *Cbr-gld-1* through reverse genetics; screening a total of two million haploid genomes , we recovered one allele, *nm68*. Mothers heterozygous for *nm41*, *nm64*, or *nm68* all produce nearly 25% sterile self progeny (23.5 ± 1.2%, n=1500, 23 ±

1.2%, n=1800, and 22.4% ± 1.1%, n =1000) respectively, consistent with these mutations as recessive loss-of-function alleles.

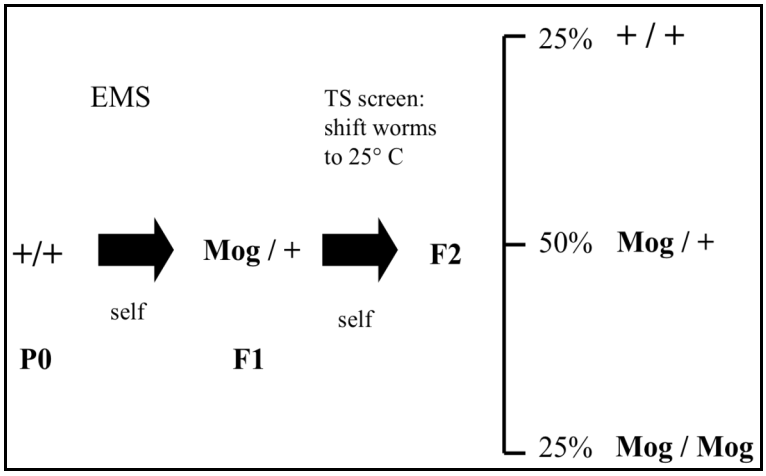


FIGURE 2. Forward mutagenesis scheme. Young adult AF16 hermaphrodites are exposed to 50mM EMS for 4 hours . F1 self-progeny are picked to fresh plates and allowed to self. Recessive alleles that masculinize the hermaphrodite germline are recovered in the F2 generation.

To initially map the *nm41* and *nm64* alleles acquired through forward mutagenesis, we tested their linkage to the candidate Mog genes *Cbr-gld-1* and *Cbr-atx-2* (Nayak et al. 2005, Maine et al. 2004). Genotyping 30 *nm41* and 30 *nm64* individual F2 steriles for a polymorphism we identified within *Cbr-atx-2* between our mapping and mutagenesis strains, we found nearly a 1:2:1 distribution of mutagensis strain to mapping strain alleles (χ = 5.84, p=0.054 for both *nm41* and *nm64*), consistent with no linkage to this gene. However, we detected strong linkage to *gld-1* by genotyping 25 *nm41* and 30 *nm64* animals and finding no recombinants between the sterile phenotype and the mapping strain allele.

Sequencing the complete *Cbr-gld-1* coding sequence in *nm41* and *nm64* sterile animals, we found a single sequence lesion in each allele: a C-to-T transition in *nm41* that transforms AA14 from glutamine to a premature stop codon, and a G-to-A transversion mutation in *nm64* that transforms AA267 from a tryptophan to a premature stop codon (Figure 3).

Finally, to show that both *nm41* and *nm64* are mutant alleles of the same locus, we

tested for their allelic complementation. We mated both *nm41*/+ hermaphrodites to *nm64*/+ males and also *nm64*/+ hermaphrodites to *nm41*/+ males. Adult hermaphrodite progeny from successful crosses were scored with light and DIC microscopy for sterility. We counted 22.1% (n=400) and 24.0% (n=300) steriles from each kind of successful cross, respectively. With DIC microscopy, 100/100 steriles scored from both experiments had phenotypes similar to those of *nm41* and *nm64* single mutants. Thus, we conclude that *nm41* and *nm64* fail to complement each other. Linkage, sequencing and complementation data taken together, we conclude that the causative mutation for both *nm41* and *nm64* lies in *Cbr-gld-1*.

The deletion allele *nm68* is a 923nt deletion in *Cbr-gld-1* beginning in exon 2 and ending in the final exon, 6. As depicted in Figure 3, *nm68* leaves 55% of the *gld-1* sequence intact, including most of the evolutionary conserved domain upstream of the KH RNA-binding domain, but removes all of the KH domain and the downstream evolutionarily conserved domain (Vernet and Artzt 1997). Based on the known and hypothesized action of GLD-1 as an RNA binding protein and the severity of this molecular lesion, we predict *nm68* to be a strong loss-of-function or a null allele.

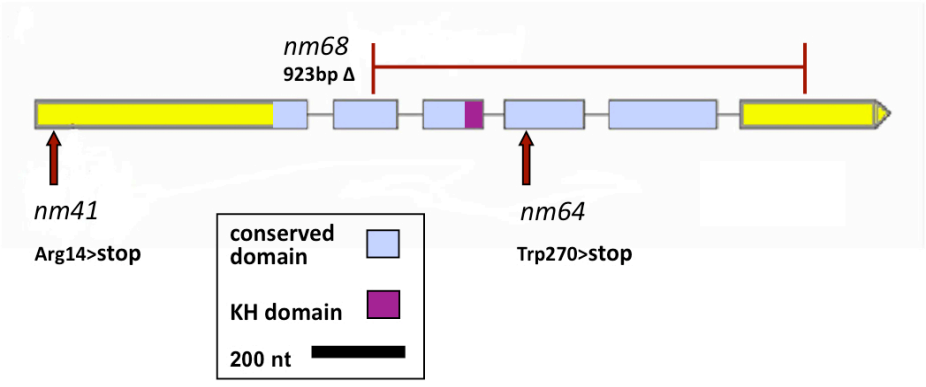


FIGURE 3. Schematic showing the location of *Cbr-gld-1* mutations. *nm41* and *nm64* are both nonsense mutations, and *nm68* is a 923 base pair deletion that removes the KH and evolutionarily conserved downstream domains of GLD-1.

We also performed a complementation test between *nm68* and *nm64* by mating *nm64*/+ XX hermaphrodites to *nm68*/+ males and scoring the progeny from successful crosses. We found 21.5% F1 steriles from 3 successful crosses (n=300 animals scored), illustrating non-complementation of the two alleles. We also examined 25 sterile XX animals with DIC and found that all 25 had phenotypes similar to *nm68* and *nm64* single mutants.

nm41, nm64, and nm68 produce mRNA, but protein product cannot be detected

We wanted to determine whether these three *Cbr-gld-1* mutations were molecular nulls. To do this, we first performed RT-PCR on homozygous *nm41*, *nm64*, and *nm68* worms using primers designed to amplify *gld-1* for each mutation. As shown in Figure 4, we found that, even for the deletion mutant *nm68*, RNA product is produced from each of the three mutated loci. Thus, the premature stop codons in the *nm41* and *nm64* alleles do not appear to induce severe nonsense-mediated decay (Cali and Anderson 1998).

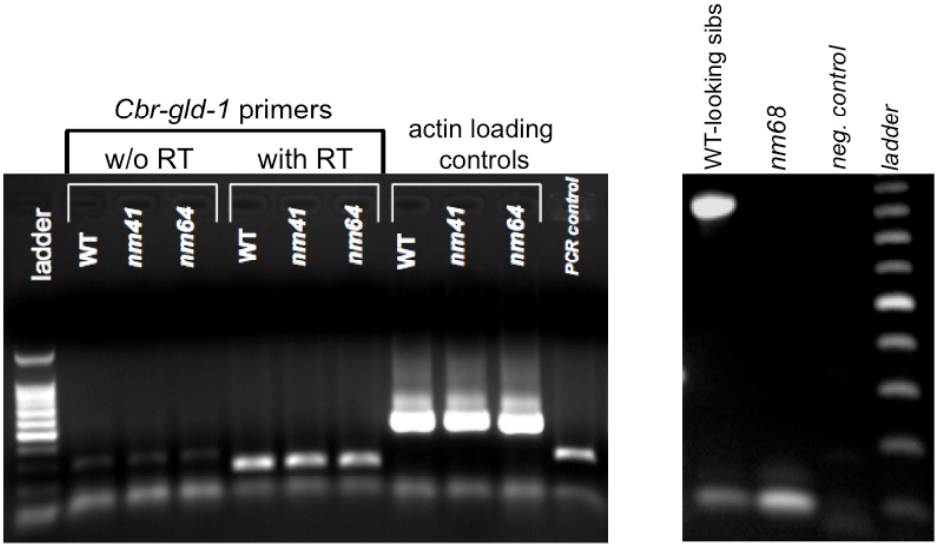


FIGURE 4. *gld-1* RT-PCR of *Cbr-gld-1* homozygous mutants . Left panel compares transcription at the *gld-1* locus between *nm41* and *nm64* and wild-type *C. briggsae* hermaphrodites. Though there is some DNA *gld-1* amplification in the no reverse transcriptase control lanes ("no RT") that contain a small intron, there is clearly transcription of *gld-1* in both *nm41* and *nm64* homozygous animals (lanes 6 and 7). Right panel compares transcription between animals

homozygous for the 923bp deletion allele *nm68* and their wild type-looking siblings (including *nm68*/+ heterozygotes). *nm68* worms produce appropriately-sized *gld-1* mRNA.

Next, we examined GLD-1 protein expression from the *nm41*, *nm64*, and *nm68* alleles by Western blot using three different anti-GLD-1 antibodies. Even though the nature of the three mutations are different from one another, and despite the fact that the three polyclonal antibodies used in this experiment recognize varied GLD-1 epitopes, we were not able to detect any GLD-1 protein in *nm41*, *nm64*, or *nm68* homozygous mutants with any antibody (Figure 5). However, we do not believe that these mutations are complete protein nulls, as their phenotypes can be distinguished from one another. Further, work on analogous mutations of *C. elegans gld-1* has revealed different molecular properties of mutations similar to these (discussed below; Francis et al. 1995a, Jones and Schedl 1995, Jones et al. 1996). As we could detect no GLD-1 protein by Western blot in the analogous *C. elegans* mutations that are known not to be nulls (Figure 5), we hypothesize that *nm41* and *nm64* may produce too little protein product to be detected by Western blot. We remain uncertain of whether *nm68* is a null mutation by this assay, but its large deletion makes it more likely.

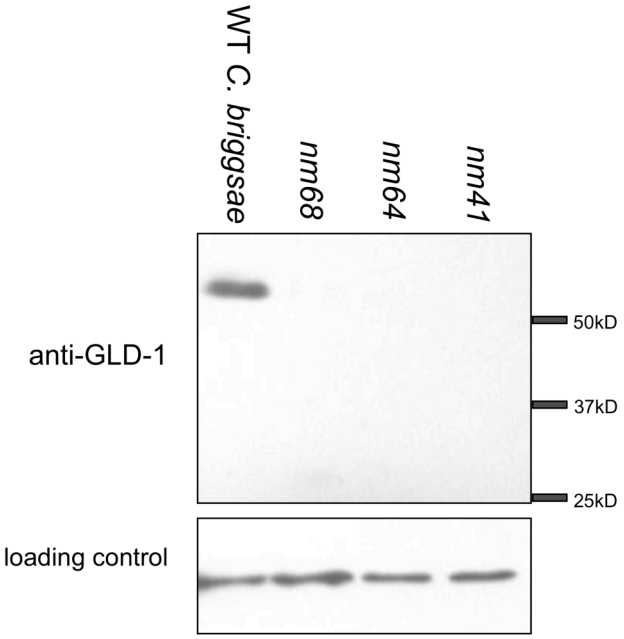


FIGURE 5. Western blot analysis of wild-type *C. briggsae* hermaphrodites and *Cbr-gld-1* mutants. 40 animals are loaded per lane. Extended gel is shown to assay for truncated proteins produced by *gld-1* mutations. Primary antibody is chicken anti-GLD-1, designed against amino acids 16-36 of Cbr-GLD-1 and is described further in dissertation Chapter 2. Tubulin is a loading control.

Cbr-gld-1(lf) alleles produce a range of mutant phenotypes, including tumorous and masculinized germlines

Mutations in *Cbr-gld-1* produce a range of germline defects in XX animals, including mitotic tumors, fully masculinized germlines, cells arrested in meiosis, germline tumors that are found only near the spermathecal-end of the gonad arm (i.e., proximal ("Pro") tumors), and various combinations of these phenotypes. No germlines possess oocytes, ooids, or cells that express the oocyte marker yolk-receptor RME-2 as judged by Western blots analysis or antibody staining of dissected gonads, except for weak staining in *nm64* (data presented below; Figures 10 and 11-14). Figure 6 illustrates the fully masculinized (left panels) and tumorous (right panels) *Cbr-gld-1* mutant phenotypes. Figure 7 categorizes the mutant germline phenotypes of *nm41* and *nm64* XX animals at both day 1 and day 3 of adulthood (in order to capture the dynamics of cell division). Phenotype data shows that though the range of germline defects is the same for all *Cbr-gld-1* mutant alleles, the prominence of each phenotype differs between alleles. For instance, *nm41* is the most likely to produce a fully sperm-filled germline and least likely to produce a full germline mitotic tumor, whereas *nm64* is the most tumorous.

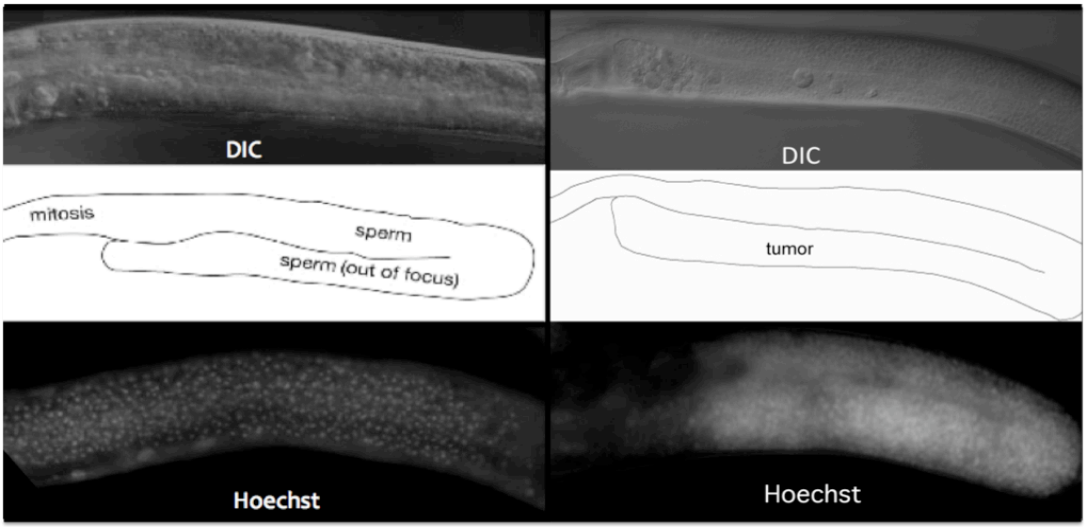


FIGURE 6. Left panels show *Cbr-gld-1* mutant masculinized germlines; right panels show *Cbr-gld-1* mutant tumorous germlines; Top panels are DIC micrographs, bottom are of lightly fixed, Hoechst-stained whole worms revealing DNA morphology. Line drawings in the middle outline gonad DIC images above. In masculinized worms (left), the compact shape of sperm (and their nuclei) are visible as small round projections in the DIC image and as small round dots in the Hoechst-stained image. In the tumorous worms (right), small, round cells fill the entire U-shaped gonad arm in the top image, and compact chromosomes in round nuclei are visible throughout the Hoechst-stained gonad below.

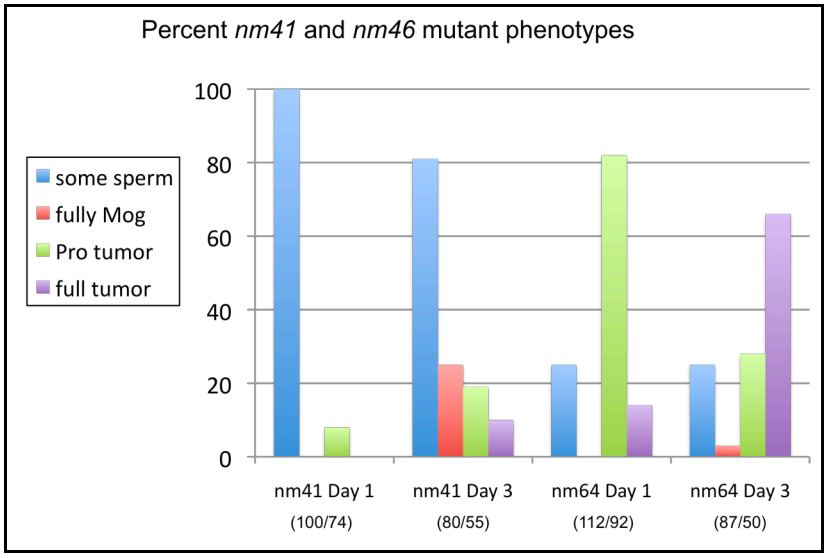


FIGURE 7. Categories of germline mutant phenotypes observed with DIC microscopy for homozygous *nm41* and *nm64* XX animals. Different animals were observed on day 1 and day 3 of adulthood. The numbers in parentheses below the X-axis present the 'number of germlines scored / number of animals', as for some animals both gonad arms were scored independently. The "Pro tumor" and "some sperm" categories are not mutually exclusive.

To determine whether *Cbr-gld-1* mutations affect males, we mated hermaphrodite and male siblings heterozygous for *nm41* or *nm64* together and looked at the male progeny from successful crosses with DIC microscopy; 25% of these animals should be homozygous *gld-1* mutants. Scoring 285 males in the *nm41* assay, we found only 8 phenotypically aberrant animals, many less than the 70 animals would one expect if *gld-1* had a penetrant phenotype in males. We then genotyped 40 of these males, including 4 of the 8 phenotypically abnormal males, to ensure that homozygous *nm41* males were still alive and among those animals scored. We detected 8/40 *nm41* homozygous genotypes, none of which belonged to the phenotypically abnormal males. We also scored 200 potential *nm64* homozygous males in the same way and found no phenotypically abnormal animals. We then genotyped 50 of these animals as above and detected 7 *nm64* homozygotes, showing that *nm64* homozygotes do not suffer lethality. Taken together, we conclude that *Cbr-gld-1* mutations have no overt phenotypic consequences in males. We note that although the above is also true in *C. elegans*, there is a minor requirement for *Ce-gld-1* for the proper number of mitotic divisions in the distal region of the male germline (Francis et al. 1995b). As we did not examine *Cbr-gld-1* male (or hermaphrodite) mutant germlines for this effect, we do not know if this is also true in *C. briggsae.*

Comparisons between C. briggsae and C. elegans gld-1 mutants

C. briggsae gld-1 mutants have defects in sex determination, tumor suppression, and meiotic progression, and these processes are also defective in the carefully-studied *C. elegans gld-1* mutants (Francis 1995a and b). Interestingly, molecular lesions identical to *nm41* and *nm64* have also been identified in *C. elegans gld-1*. As depicted in Figure 8, *Ce-gld-1* mutations *q286* and *q395* are the same nonsense mutation as *nm64* (Jones and Schedl 1995). Further, the *C. elegans gld-1* allele *q93oz50* was isolated in a screen to suppress the *Ce-gld-1(q93)* mutant

phenotype; the *oz50* intragenic suppressor is the same nonsense mutation as *nm41* (Francis et al. 1995a, Jones and Schedl 1995; Figure 8). The only null allele of *C. elegans gld-1*, *q485*, results from an 82 base pair deletion early in *gld-1*'s protein coding region that produces no mRNA transcript or protein product (Jones and Schedl 1995, Jones et al. 1996; Figure 8).

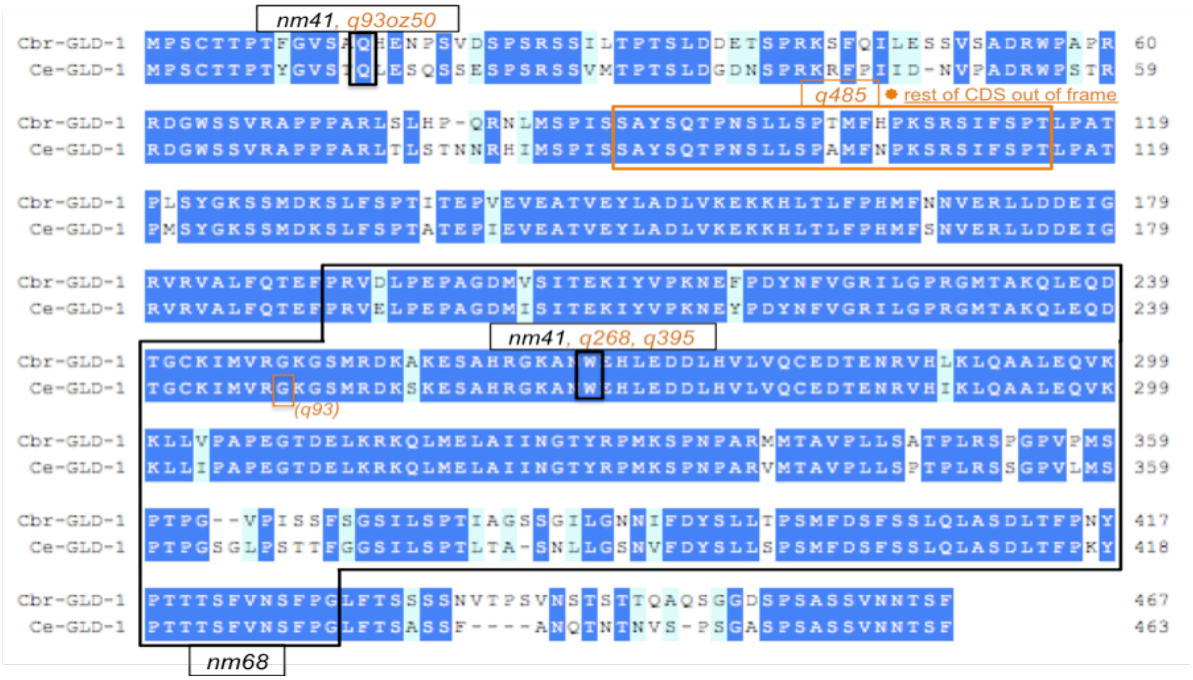


FIGURE 8. An amino acid alignment of Cbr-GLD-1 and Ce-GLD-1. *Cbr-gld-1* mutations are indicated with black boxes and font (*nm41, nm64*, and *nm68*), and the corresponding *Ce-gld-1* mutations are in orange font (as well as the *q93* mutation that contributes to *q93oz50*); the *Ce-gld-1* null allele *q485* is also shown. Lighter blue shading is for conservative amino acid substitutions, and dark blue designates identical amino acids; no shading indicates more radical amino acid substitutions and insertions/deletions.

Despite the strong sequence conservation between *C. elegans* and *C. briggsae* GLD-1 and similar molecular lesions in the two species, there are intriguing differences in *Ce-gld-1* and *Cbr-gld-1* mutant phenotypes. Most importantly for this study, when gonads from homozygous XX *C. elegans gld-1* mutants with genotypes *q485*, *q286* or *q93oz50* are dissected and stained with an antibody against the oocyte yolk receptor RME-2, tumorous and meiotic arrested cells

stain strongly with this oocyte marker, but do not stain with an antibody to the Major Sperm Protein (Lee and Schedl 2001, this work). However, as Nayak et al. 2005 first indicated by inhibiting *gld-1* with RNA interference (RNAi), loss of *gld-1* surprisingly produces the opposite phenotype in *C. briggsae*, masculinized germlines. We confirmed opposite major sex determination roles of *gld-1* in the two species in a direct comparison of the *C. briggsae* deletion mutant *nm68* and nonsense mutants *nm64* and *nm41* with the *C. elegans gld-1* null mutant *q485* and corresponding *C. elegans* mutations by Western blot (Figure 9). This is especially important because as both species *gld-1* mutants display gross cell-cycle/meiosis defects, it is not possible to assign all cells in all animals a germline sexual fate by gamete morphology alone. Probing with an antibody to the oocyte yolk receptor RME-2, a marker of oocyte fate, and an antibody to Major Sperm Protein (MSP), a marker for sperm fate, we find that all the *C. elegans* mutants produce RME-2, despite their tumorous or meiotic-arrested germlines. However, *Cbr-gld-1 nm68* and *nm41* only produce MSP. This is despite the fact that *rme-2* is a direct mRNA target of GLD-1 translational repression both in *C. elegans* and in *C. briggsae*, such that one might expect its protein to be easily found in *Cbr-gld-1* mutant germlines (Lee and Schedl 2001, Nayak et al. 2005, dissertation Chapter 2). Interestingly, *C. briggsae nm64* shows a weak RME-2 signal, and this is confirmed by antibody staining of dissected gonads (Figures 12 and 13). Francis et al. 1995a showed that the feminizing character of the corresponding *C. elegans* mutations *q268* and *q395* was due to a gain-of-function property; this mutational property may be conserved in *C. briggsae*.

Importantly, we note that RME-2 is only a single marker of oocyte cell fate, and thus we cannot say assuredly that *Cbr-gld-1* mutant germ cells do not have oocyte character. Our work does show, however, that these germ cells at least do not have the same oocyte character as *C. elegans gld-1* mutant germ cells. These results and phenotype data taken together demonstrate

that while *C. elegans gld-1*(lf) mutations cause feminization of the germline, *C. briggsae gld-1* mutants instead suffer from germline masculinization.

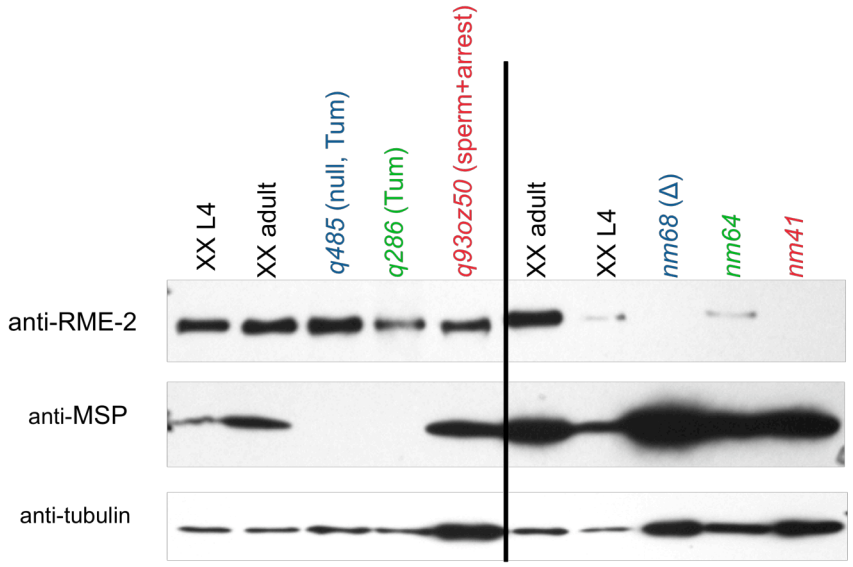


FIGURE 9. Western blot analysis of wild-type *C. elegans* and *C. briggsae* hermaphrodites and XX animals homozygous for the *C. elegans gld-1* null allele *q485*, *C. briggsae* deletion allele *nm68*, *C. briggsae* nonsense mutations *nm41* and *nm64*, and the corresponding mutations in *C. elegans q268* and *q93oz50. C. elegans q485* and *q268* mutants possess a germline tumor ("Tum"), and the *q93oz50* mutant phenotype is sperm and cells in meiotic arrest. *C. briggsae* mutants are as described in text. Protein is probed with an antibody to the oocyte yolk receptor RME-2 and an antibody to the Major Sperm Protein, MSP, in order to assess the sexual fate of mutant germlines in the absence of true gametogenesis. The anti-MSP staining in *Cbr-gld-1* mutants likely only derives, however, from mature sperm and the few recognizable spermatocytes, not from grossly tumorous cells (Figures 10 and 11). Tubulin is a loading control; *C. briggsae gld-1* mutants reproducibly produce more tubulin than the same number of wild-type animals.

We also investigated other sex determination parallels between the *Cbr-gld-1* and *Ce-gld-1* mutations. First, the feminizing phenotype of *C. elegans gld-1* was revealed in part by testing for its haploinsufficiency in sex determination. Screening *Ce-gld-1*(*q485* null)/+ XX animals, Francis et al. 1995a found that 2% possessed the stacking oocyte feminized germline phenotype, and animals that did produce sperm had a smaller brood size than wild-type, which is evidence of sperm limitation. Further showing that this property was due to a loss of *gld-1*

function in the *q485* null mutant, Francis et al. concluded that *Ce-gld-1* is haploinsufficient for hermaphrodite spermatogenesis. We thought it possible that even though the major sex determination function of *C. briggsae gld-1* is to promote hermaphrodite oogenesis, it might also have a conserved minor role in promoting hermaphrodite spermatogenesis, and this role might be revealed through its haploinsufficiency, as it is in *C. elegans*. Conversely, perhaps *C. briggsae gld-1* is haploinsufficient for the same sex determination function revealed by its homozygous mutants, that is, oocyte production. If *Cbr-gld-1* were haploinsufficient for specifying the oocyte fate, then *gld-1*/+ mothers would overproduce sperm at the expense of oocytes and thus would have a larger brood size than wild-type animals (as sperm, not oocytes, are the limiting factor in an individual hermaphrodite's fertility).

To test for these two different kinds of *Cbr-gld-1* haploinsufficiencies, we let *nm41*/+, *nm64*/+, and *nm68*/+ mothers self and then scored their progeny for both the presence of rare feminized animals and for total brood size. Scoring 1000 progeny from 10 heterozygous mothers of each genotype, ½ of which should be heterozygous for a *gld-1* allele, we found less than 1% feminized progeny for all three experiments. While not zero, this low percentage is not convincingly different from the number of 'naturally feminized' animals one might find in *C. briggsae* wild-type stocks. Thus, we found no strong evidence for haploinsufficiency of sperm production.

Additionally, comparing the average brood sizes for *nm41*/+ and *nm64*/+ mothers to wild-type animals of the same genetic background (that is, the sisters of these animals that only produce wild-type progeny), we find no statistically significant differences in the average brood sizes between 30 heterozygous mothers and 30 wild-type mothers in each experiment (data not shown). Thus, we cannot find evidence of a penetrant *Cbr-gld-1* haploinsufficiency for hermaphrodite oocyte production with *nm41* or *nm64*.

Questionable sexual identity of the Cbr-gld-1 germline tumors

In *C. elegans*, GLD-1 is known to be important for four cellular processes in the adult germline: a redundant, minor role in the decision of germ cells to enter meiosis versus staying in mitosis in the distal germline, a role in the specification of hermaphrodite sperm, a role in proper oocyte formation, and a major role in sustaining the progression of cells in meiotic oogenesis (Francis et al 1995a, Francis at al 1995b, Jan et al. 1999; see Hansen and Schedl 2006 and Kimble and Crittenden for reviews of mitosis). *gld-1's* role in keeping oocyte-fated cells committed to meiosis is dramatically demonstrated by the tumor formation of some *C. elegans gld-1* homozygotes, including the null allele *q485*. In *C. elegans*, Francis et al. showed that by investigating germ cell DNA morphology over time with DAPI staining, these *gld-1* tumors result from meiotic cells re-entering mitosis, and only when the germline is set to produce oocytes, not sperm, does this germline tumor form (1995a and 1995b).

Interestingly, when forced into female mode through loss of the sperm-promoting *fog-3* gene product (Ellis and Kimble 1995, Chen et al. 2000), Nayak et al. 2005 showed that *C. briggsae* double *fog-3 gld-1* RNAi germlines also produce a feminized tumor that stains with an antibody against the oocyte yolk receptor RME-2, just as in *C. elegans*. Thus, the role of *gld-1* in keeping oocyte-fated cells committed to meiosis is conserved between *C. elegans* and *C. briggsae*; indeed we find that this role is conserved widely in the genus *Caenorhabditis* (data presented below).

However, unlike the tumor that forms in *C. briggsae* when germlines are forced into the female fate, single *Cbr-gld-1* presumed-null homozygotes do not produce cells that express RME-2, nor do they have cells with morphological oocyte character. We confirmed this finding with antibody staining of dissected gonads, finding that *C. briggsae gld-1* mutant *nm68* and *nm41* germlines do not possess any RME-2 (Figures 10 and 11); *nm64*, as mentioned above,

demonstrates weak staining that may be a conserved gain-of-function character with the analogous *C. elegans* mutant *q268* (Figure 9; Figures 12 and 13). A comparison of stained gonads between animals homozygous for the *C. briggsae* putative null allele *nm68* (Figure 10) and the *C. elegans* null allele *q485* (Figure 14) clearly illustrates this opposite major sex determination mutant phenotype in the two species.

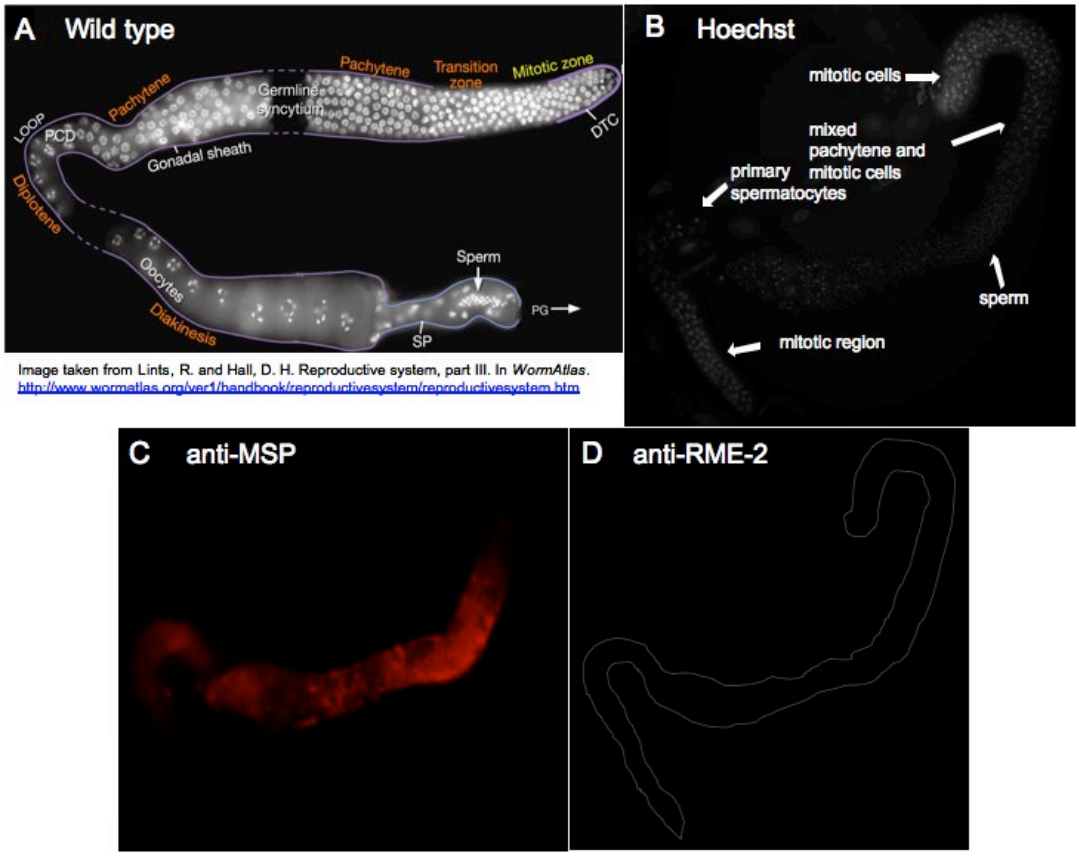


FIGURE 10. *Cbr-gld-1(nm68)* dissected gonads stain with an antibody to Major Sperm Protein. (A) Image from Lints and Hall (2009). DAPI-stained dissected gonad from a wild-type *C. elegans* hermaphrodite to illustrate *Caenorhabditis* germline chromosome morphology. Image indicates the progression of germ cells from the self-renewing germ cell source, the "mitotic zone," to fully differentiated gametes (oocytes and sperm); (B) dissected *nm68* gonad arm stained with Hoechst-33258 to identify chromosome morphology. Though the mitotic region and transition zone are normal, the worm overproduces sperm compared to wild-type animals and has spermatocytes even in adulthood. At the proximal end of the gonad is ectopic germline proliferation, and just distal to the tumor is a region of mixed pachytene-like cells and mitotic cells; (C) same gonad arm as in (B), stained with anti-MSP antibody, a marker for the sperm fate. anti-MSP stains primary spermatocytes and sperm, but not mitotic cells; (D) same gonad arm as in (C), stained with an antibody to the oocyte yolk receptor RME-2 as a marker for oocyte fate. anti-RME-2-stained image is overexposed to detect potential weak staining.

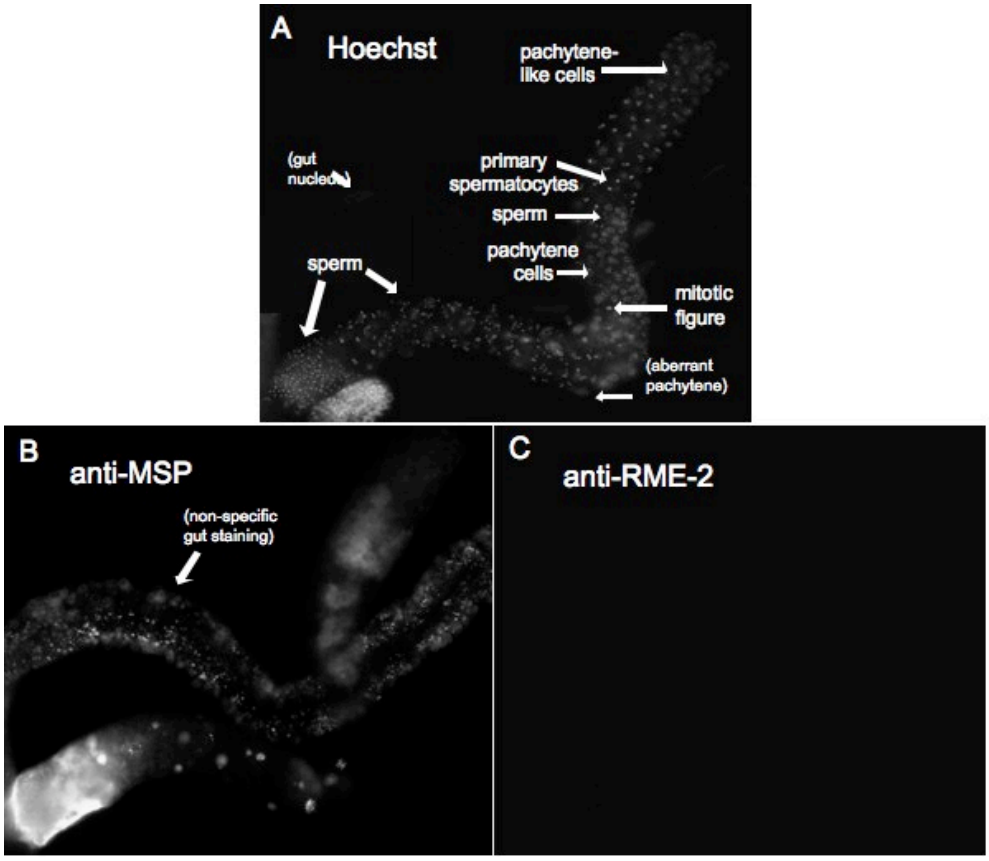


FIGURE 11. *Cbr-gld-1(nm41)* dissected gonads stain with an antibody to Major Sperm Protein. (A) Hoechst-stained *nm41* dissected gonad to reveal chromosome morphology. Only cells proximal to the transition zone are shown. This adult gonad still contains primary spermatocytes and overproduces sperm. Gonad also has a region of mixed pachytene and mitotic cells; (B) and (C) as in Figure 11. anti-MSP antibody stains mature sperm and primary spermatocytes. anti-RME-2-stained image is overexposed to detect potential weak staining.

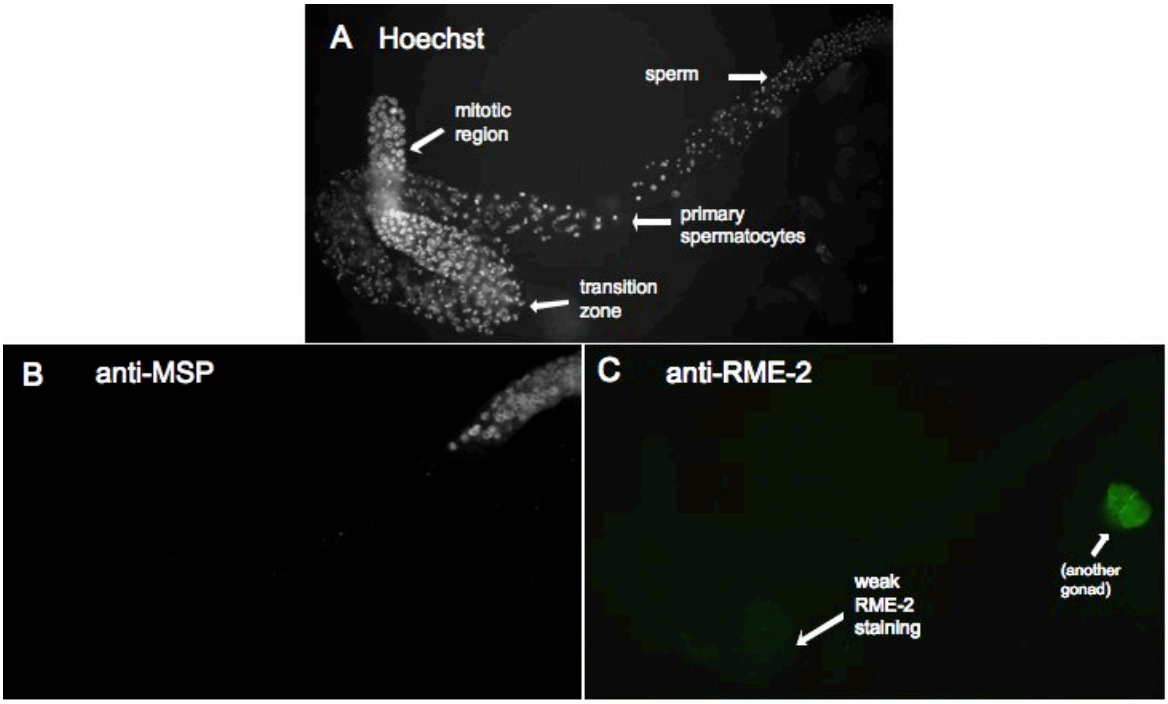


FIGURE 12. *Cbr-gld-1*(nm64) dissected gonads display both MSP and weak RME-2 staining. (A) Hoechst-stained *nm64* dissected gonad arm showing a normal mitotic region but an abnormally large transition zone and cells with unusual DNA morphology. As in Figures 10 and 11, this gonad arm has an over-proliferation of sperm compared to wild-type animals and still contains spermatocytes; (B) and (C) as in Figures 10 and 11. (B) anti-MSP antibody stains only the mature sperm; (C) *nm64* gonad arms sometimes exhibit weak RME-2 staining in meiotic-like cells near the transition zone. anti-RME-2-stained image is overexposed to detect potential weak staining.

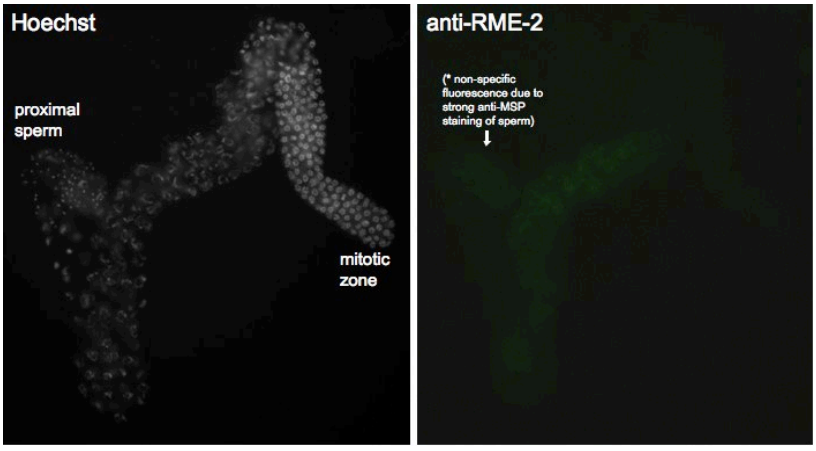


FIGURE 13. *nm64* dissected gonad stained with Hoechst (left) and anti-RME-2 antibody (right). RME-2 is detected in cells proximal to the transition zone with abnormal chromosome morphology. anti-RME-2-stained image is overexposed to detect potential weak staining.

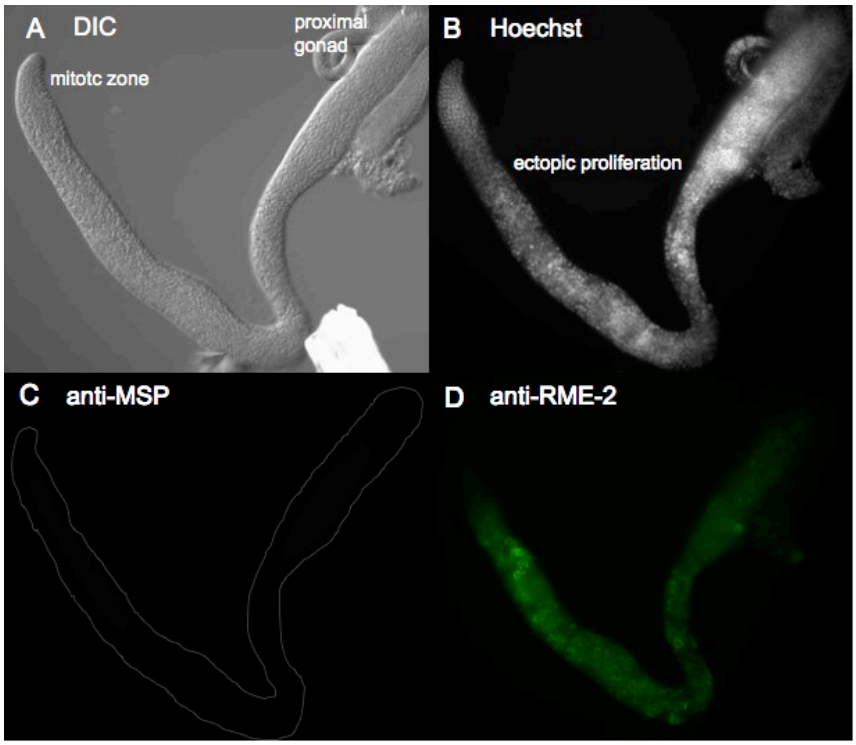


FIGURE 14. *C. elegans gld-1*(q485) null dissected gonad arm stained with (B) Hoechst dye for chromosome morphology, (C) anti-MSP antibody to mark the sperm fate, and (D) anti-RME-2 antibody to mark the oocyte fate.

Despite these opposite sexual phenotypes, however, the formation of germline tumors in *gld-1* mutant germlines appears similar in *C. briggsae* and *C. elegans*. Examining the Hoechst-stained DNA morphology of *C. briggsae* mutant germlines, Figures 10 and 11 suggest that *gld-1* tumors seem to result from a failure of cells that have entered meiosis to remain in meiosis, as in *C. elegans*. This is indicated by two observations, that tumor formation only begins proximal to the transition zone, where wild-type germlines first enter meiosis and beyond where mitotic cells are ever found, and that by Hoechst-staining of chromosome morphology, we find mitotic cells (with condensed nuclei or mitotic metaphase and anaphase figures) mixed with cells that display the characteristic pachytene, synapsed DNA.

Thus, both *C. elegans* and *C. briggsae gld-1*(null or presumed null) mutant germlines can form a tumor that appears to result from meiotic germ cells aberrantly re-entering mitosis.

Unlike in *C. elegans* however, we cannot determine a sexual fate for the *C. briggsae* tumorous cells themselves, as they do not stain with antibodies to MSP or RME-2. Therefore, four explanations for the origin of these tumorous cells seem possible: 1) they undergo meiosis fated to become sperm, and still retain sperm character upon re-entry into mitosis (analogous to the RME-2-expressing tumor cells in *C. elegans*), but do not express MSP in particular; 2) they undergo meiosis fated to become sperm, but subsequently fully de-differentiate and lose sexual identity; 3) they are actually female-fated cells that cannot progress through the oocyte cell cycle but also do not express RME-2; or 4) the cells are sexually uncommitted at the time of their aberrant re-entry into mitosis.

To investigate a potential spermatogenic origin of *Cbr-gld-1* tumors, we constructed double mutants using a strong loss-of-function mutation in *Cbr-tra-2, nm1* (Kelleher et al. 2008) and all three *Cbr-gld-1* mutant alleles. Loss-of-function mutations in *tra-2* masculinize both the germline and soma (incompletely) of XX animals, causing them to produce only sperm as in a wild-type males. We performed single worm matings between *gld-1*(lf)/+ hermaphrodites and *nm1*/+ males to obtain double heterozygous F1 progeny that would yield both tumorous/masculinized and pseudomale offspring. We then examined the germlines of 60 F2 *tra-2* pseudomales for each experiment with DIC microscopy, 1/4 of which should be homozygous for either *Cbr-gld-1*(nm41), (nm64), or (nm68). We find that 100% of the animals examined have a typical *tra-2* male germline, producing only sperm and no tumors. Additionally, we genotyped at least 30 of these pseudomales animals in each experiment and found that double *tra-2; gld-1* mutants were indeed among the animals examined. Thus, we infer that either *gld-1*(lf) germline tumors are not developed from spermatogenic cells, or that instead *tra-2* XX pseudomales are masculinized enough to have a *gld-1* mutant response like XO males –that is, normal spermatogenesis.

To further investigate the sexual fate of *C. briggsae gld-1* mutant germline meiotic cells and tumors, we attempted to identify sperm and oocyte "marker" transcripts that we could use in RT-PCR or *in situ* hybridization to sex the tumorous cells. We used the *C. elegans* literature and WormBook (e.g., L'Hernault 2006) and also data from whole genome microarray experiments (including Reinke at al. 2004) to find genes that were most enriched in either *C. elegans* male or hermaphrodite germlines. We then identified a *C. briggsae* homolog or ortholog of these genes on WormBase by BLAST, and Dorothy Johnson, an undergraduate working in our laboratory, designed primers to amplify these genes from *C. briggsae* wild-type males and mutant *nm38* female RNA (dissertation Chapter 3).

Unfortunately, surveying about 30 genes, we were not able to identify any transcripts that were sex-specific in both *C. elegans* and *C. briggsae*. Table 1 contains some of the *C. elegans* genes we assayed, their *C. briggsae* ortholog/homolog, and the qualitative status of our RT-PCR assay in *C. briggsae* wild-type males. We note that many "oogenesis enriched" genes, like *egg-1/2*, *mrp-4*, *pos*-1, and *spn-4* are likely to function in the male soma or germline, and we indeed find expression by RT-PCR of these genes in wild-type *C. briggsae* XO males. We also, however, detected expression of transcripts thought to be specific to *C. elegans* hermaphrodites in wild-type *C. briggsae* males: *oma-1/2*, *cpg-2*, and *cey-2/3*. We note that none of the "oogenesis enriched" genes in Table 1 are on the X chromosome in *C. elegans*, and thus their transcription is possible in the male germline (Kelly et al. 2002). (Interestingly, we also find ten genes among the top 50 most female-biased in the Reinke et al. 2004 "oogenesis enriched" microarray gene set have no homologs in *C. briggsae*, compared to the 95% overall homology frequency between *C. elegans* and *C. briggsae* (Hillier et al. 2007).)

Table 1

C. elegans gene	*C. briggsae* ortholog	*C. briggsae* male expression
cey-2/3	homolog: WBGene00035267	present in males
cpg-2	homolog: WBGene00039115	present in males
C05C10.5	WBGene00030729	present in males
C17E4.3	WBGene00026537	present in males
egg-1/2	homolog: WBGene00030306	present in males
mex-3	WBGene00027635	present in males
mrp-4	WBGene00025084	present in males
oma-1/2	homolog: WBGene00040414	present in males
pie-1	WBGene00032830	present in males
pos-1	WBGene00032403	present in males
spn-4	WBGene00038303	present in males
W05F2.3	WBGene00035333	present in males
zif-1	WBGene00029058	present in males
F07A11.2	WBGene00025820	present in males
Y39A1A.12	WBGene00037778	present in males
C01G8.1	WBGene00033734	present in males
W01A11.2	WBGene00038245	present in males

Considering spermatogenesis-enriched genes, we were also unable to identify transcripts that were male-specific in both *C. briggsae* and *C. elegans* in order to compare side-by-side their transcription in *Ce-gld-1* vs. *Cbr-gld-1* tumorous germlines. We found that *Cbr-spe-26* and *Cbe-spe-4* transcription was male-specific in *C. briggsae* wild-type males compared to mutant *Cbr-fog(nm38)* females. However, surprisingly, neither gene was male-specific in *C. elegans*, as we detected transcription in *Ce-fog-2* female mutants (data not shown). Other genes like *Ce-Y69E1A.1* were expressed in *C. elegans* males and not in *fog-2* mutant females, but we detected transcription in *C. briggsae nm38* females. As RT-PCR is a sensitive technique, these patterns may be due to technical variability between experiments, or they might reveal true gene expression differences between mutant *Ce-fog-2* and *Cbr-nm38* mutant females. Taken together, however, our investigation of female-specific vs. male-specific transcripts in *C. elegans*

and *C. briggsae* did not provide us with a way to determine the sex of *Cbr-gld-1* mutant meiotic cells or tumors.

Cbr-tra-1(nm2) is epistatic to Cbr-gld-1

To see where *gld-1* acts in the *C. briggsae* germline sex determination pathway to promote hermaphrodite oogenesis, we desired to performed epistasis analysis with *Cbr-gld-1* mutants and other strong loss-of-function or putative null *C. briggsae* sex determination alleles. These experiments were hindered in two ways, first by the variability of phenotypes in homozygous *gld-1* mutant animals (see above), and second by the general dearth of phenotypic markers and absence of balancer chromosomes in *C. briggsae.*

Though different mutations in *C. elegans* genes have been used to pseudo-balance *Ce-gld-1*, *gld-1*'s chromosomal location is different in the two species: it is found in the middle of chromosome I in *C. elegans* but is at the left end of chromosome I in *C. briggsae*, and micro-synteny around *gld-1* is only partially conserved in the two species (unpublished observations). To aid our *Cbr-gld-1* investigations, we attempted to derive a phenotypic marker linked to *gld-1* through deletion screening for homologs of *C. elegans* genes lying ~500,000bp (i.e., 3 map units at this end of chromosome I; Hillier et al. 2007) in either direction of *Cbr-gld-1* that might produce a phenotypic effect when mutated (like a *dpy* or *unc* gene). Using WormBase to scan this region of the *C. briggsae* genome on a gene-by-gene basis, we identified 3 genes, CBG11865, CBG00294, and CBG11852, which met these criteria. We attempted to obtain deletion alleles of these genes in a reverse genetic screen conducted by Robin Hill in our laboratory. From this work, we detected a deletion mutation in CBG11865, but could not isolate it through successive generations of worm husbandry. Thus, thus *Cbr-gld-1* remains unmarked.

Because homozygous loss-of-function mutations in *Cbr-tra-1* masculinize the soma, and

thus help to mark potential *tra-1; gld-1* double mutants, we proceeded with double mutant analysis for this gene. *Cbr-tra-1* is generally a female-promoting gene that acts near the end of the sex determination pathway in both *C. briggsae* and *C. elegans*. *Cbr*-tra-*1(nm2)* is a hypomorphic mutation that completely masculinizes the XX soma and, as in true in other *Ce-tra-1* and *Cbr-tra-1* mutants, causes both sperm and oocytes to be made in XX and XO homozygous mutants (Kelleher et al. 2008). We sought to know whether *gld-1(nm68)* or *gld-1(nm41)* could suppress *nm2* oocyte formation. In this experiment, we mated homozygous *nm2* XX pseudomales to *gld-1(lf)*/+ hermaphrodites at 20°C. These pseudomales have limited cross-fertility before their mutant germlines switch to producing oocytes. F1 L4 hermaphrodite cross-progeny were singled and selfed at 25°C. The phenotypes of the F2 generation from doubly heterozygous F1 mothers (which produce both pseudomale and tumorous/masculinized self-progeny) were first counted, and then potential double mutant animals (i.e., those with a Tra, masculinized soma) were scored with Nomarski optics for germline sex and status of germ cell progression.

From the F2 generation, we counted 25.7 ± 2.3% and 23.5 ± 1.3% pseudomales, and 17.3 ± 0.3% and 17.0 ± 1.2% tumorous/masculinized animals produced by *nm68/+; nm2/+* and *nm41/+; nm2/+* mothers, respectively, which is as expected if *gld-1; tra-1* double mutants possess a pseudomale soma. Then we examined 150 pseudomales with Nomarski optics from the *gld-1(nm68); tra-1(nm2)* experiment (1/4 of which should be *nm68* homozygotes) and 90 pseudomales from the *gld-1(nm41)*; *tra-1(nm2)* experiment (1/4 of which should be *nm41* homozygotes). As shown in Table 2, we find that the breakdown of germline phenotypes is largely consistent with the phenotypes of *tra-1(nm2)* mutants alone (Hill and Haag 2009). Thus, although we have not phenotyped animals with a known genotype in this experiment, we

conclude that *gld-1*(lf) mutations cannot suppress the oocyte production in *tra-1(nm2)* animals; in fact, they seem to augment it.

Table 2	**Hill and Haag, 2009 *nm2* phenotypes**	***nm2;* ¼ *nm68***	***nm2;* ¼ *nm41***
oocytes/ooids	35%	42.5%	60.6%
only sperm	55%	50%	36.4%
abnormal somatic gonad/germ cells	10%	7.5%	3%

Cbr-gld-1(nm41); Cbr-fem-3(nm63) double mutants are self-fertile

While testing different sex determination double mutant combinations with *Cbr-gld-1* in epistasis analysis, we fortuitously came upon a pair that resulted in self-fertility: *Cbr-gld-1(nm41)*; *Cbr-fem-3(nm63)*.

Cbr-fem-3*(nm63)* is a large deletion and presumed null allele of *Cbr-fem-3*. Though loss of function alleles of *C. elegans fem-3* result in complete feminization of both XX and XO animals, in *C. briggsae*, *nm63* has no obvious phenotype in XX animals, though it transforms XO males into self-fertile hermaphrodites (Hill et al. 2006). We created *gld-1; fem-3(nm63)* double mutants with two *gld-1* alleles, the large deletion allele *nm68* and nonsense mutation *nm41*.

To determine the *gld-1(nm41)*; *fem-3(nm63)* double mutant phenotype, we mated *nm41*/+ hermaphrodites to *nm63*/+ males and selfed hermaphrodite L4 progeny from successful crosses at 25°C. Surprisingly, double heterozygous F1 mothers produced only 16.5 +/- 0.8% steriles (n=132, from 8 mothers) in the F2 instead of the 22.3 +/- 1.5% steriles produced by *nm41*/+; +/+ mothers in the same experiment (n=178, from 8 mothers). By unpaired t-test, this is a statistically significant difference (p=0.02). As *gld-1*-like sterility was the only mutant phenotype observed on the double mutant plates, this suggested that perhaps *nm41; nm63* double mutant animals were self-fertile. To verify this, we genotyped 94 self-fertile animals

from double heterozygous F1 mothers for *nm41* and found 5 *nm41* homozygotes (close to the expected value 1/13 = ~7 animals). Thus we conclude that *fem-3*(nm63) can "rescue" *Cbr-gld-1*(nm41) to self-fertility.

However, this "rescuing" ability of *fem-3*(nm63) is limited to the weaker *Cbr-gld-1* allele, *nm41*. Performing a similar experiment with the deletion allele *gld-1*(nm68), we find that the percentage of sterile animals produced by *nm68*/+; *nm63*/+ mothers and *nm68*/+ mothers are not statistically different from one another (p=0.18). We thus conclude that *nm68; nm63* double mutants are sterile, just as *nm68* single mutants are, though we do not know the sexual fate of the undifferentiated/tumorous cells in the double mutant germline.

Cbr-gld-1 can rescue the C. elegans gld-1(null) allele q485

C. briggsae GLD-1 and *C. elegans* GLD-1 have very high sequence identity, almost 90% at the amino acid level. This high sequence similarity, as well as their common roles in directing oocyte progression and repressing the translation of RME-2 as shown by Nayak et al. (2005), suggest that GLD-1 in both species has the same biochemical activity and pattern of action. We therefore can ask: why is the major sex determination mutant phenotype of *gld-1* feminizing in *C. elegans* vs. masculinizing in *C. briggsae*? One hypothesis is that a difference(s) in the messenger RNA targets and/or protein binding partners of *C. elegans* GLD-1 vs. *C. briggsae* GLD-1 might be responsible for their different major sex determination functions, as opposed to a change in GLD-1 protein function itself.

To formally demonstrate that GLD-1's different sex determination role in *C. briggsae* and *C. elegans* is not likely due to a change in the GLD-1 protein, we sought to rescue the *C. elegans gld-1* null allele *q485* with wild-type *Cbr-gld-1*, driven with *C. briggsae* 5' and 3' regulatory sequences. We bombarded construct pAD-g6 into *Ce-unc-119*(ed3) worms (Praitis

2001) and recovered animals that were both wild-type-moving (i.e., *unc-119*-rescued) and had bright GFP expression in their pharynxes. In this way, we recovered 12 potentially integrated lines.

To test for *Cbr-gld-1* rescue of *C. elegans gld-1* null homozygotes, we first mated pAD-g6-containing hermaphrodites to wild-type *C. elegans* males for each the 12 potentially integrated lines. We then mated transgenic male F1 cross-progeny to pseudo-balanced hermaphrodites of the strain BS3156 [*unc-13*(e51) *gld-1*(q485)/*dpy-18*(h662) I]. From this cross, half of the pAD-g6 transgenic progeny (identified by GFP expression in their pharynx) will be *unc-13*(e51) *gld-1*(q485)/+. We then singled and selfed five GFP-expressing, wild-type-looking hermaphrodite L4 cross-progeny to look for fertile (i.e., rescued) Uncs (i.e., *gld-1*(q485) homozygotes) in the next generation. In parallel, we also performed a control cross by mating *C. elegans* wild-type males into BS3156 heterozygotes, selfing hermaphrodite L4 progeny from successful crosses, and counting the number of self-fertile Uncs in the F2. In this control cross, we recovered <2% self-fertile Uncs (n=400), presumably the result of recombination between *unc-13* and *gld-1*.

For the 12 different experimental rescue crosses, we found 2 lines, CP113 and CP114, that produced self-fertile, pAD-g6-containing *Ce-gld-1(q485)* homozygotes with high penetrance. Figure 15 illustrates this rescue, in which sperm and oocyte production, as well as rachis formation, are restored.

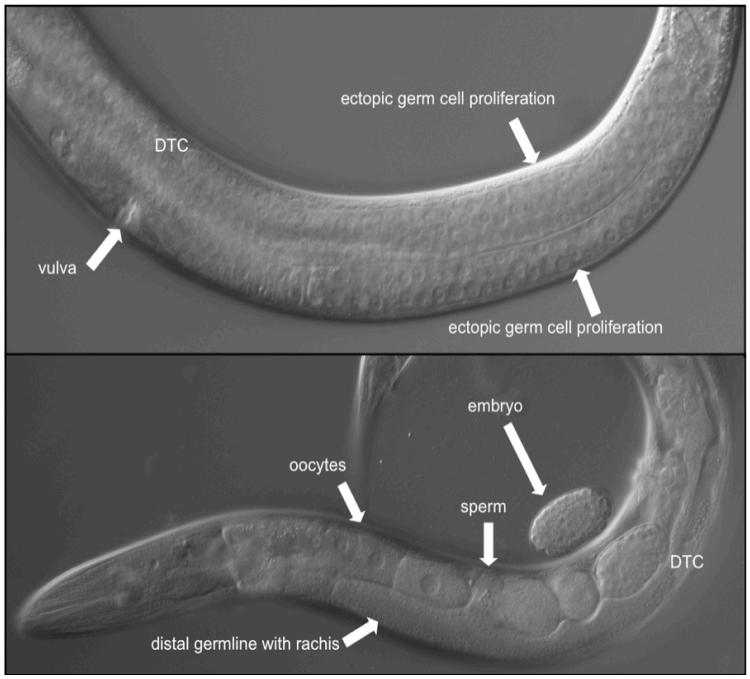


FIGURE 15. *Ce-gld-1*(q485 null) homozygotes are rescued by *C. briggsae gld-1*. Top panel depicts a *C. elegans gld-1* null XX animal, with ectopic germ cell proliferation and no sperm. Bottom panel depicts a *Ce-gld-1* null animal rescued to self-fertility by pAD-g6, containing the *Cbr-gld-1* locus. All obvious *q485* defects have been rescued, including oogenic meiotic progression, specification of hermaphrodite sperm, and formation of the germline rachis. (DTC is the distal tip cell of the somatic gonad.)

We confirmed expression of pAD-g6 in the two *Ce-gld-1* rescued strains by Western blot (Figure 16). We found that both CP113 and CP114 express GLD-1 and its HA protein tag, confirming Cbr-GLD-1 expression in rescued *C. elegans* animals. Additionally, we used single-worm PCR to show that rescued, self-fertile CP113 and CP114 animals are indeed homozygous for the *q485* mutation (Figure 17).

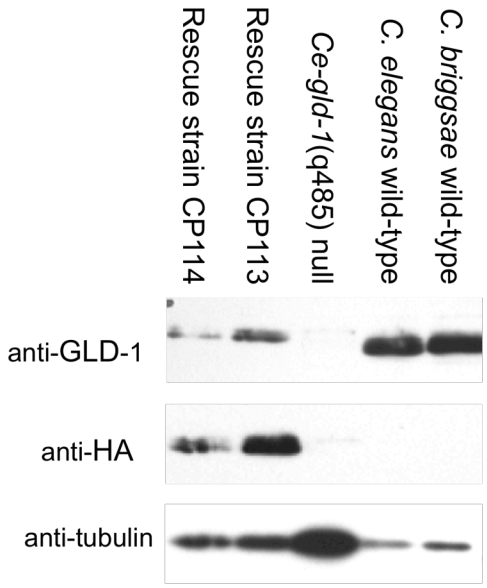


FIGURE 16. Western blot showing both GLD-1 and HA expression in the two *Ce-gld-1*(q485 null) rescued strains CP113 and CP114. Wild type *C. elegans* and *C. briggsae* hermaphrodites have abundant GLD-1 but no HA protein, and *Ce-gld-1* null homozygotes have neither GLD-1 nor HA. Blots were first probed with rabbit anti-GLD-1, then stripped and reprobed with anti-tubulin (loading control), then stripped again and probed with anti-HA. Each lane contains 40 animals.

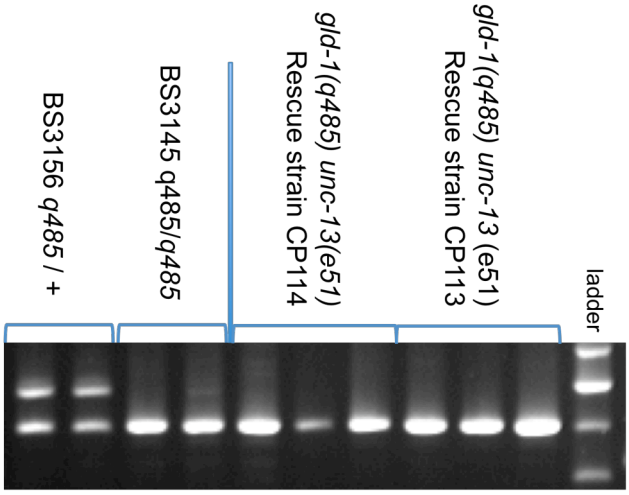


FIGURE 17. Confirmation that self-fertile *unc-13* homozygous animals are *Ce-gld-1*(q485) homozygotes in CP113 and CP114. The *q*485 allele is an 82bp deletion at the *gld-1* locus. *Ce-gld-1* was amplified from *q485* heterozygotes and homozygotes and from Uncs of the strain BS3156 rescued to self-fertility in CP113 or CP114.

We also identified a small percentage of CP113 and CP114 animals with feminized germlines. These animals may be partially rescued *Ce-gld-1* mutants that express enough Cbr-GLD-1 to overcome the oogenesis tumor mutant phenotype but not enough GLD-1 to initiate hermaphrodite spermatogenesis. Alternatively, as *gld-1* in *C. elegans* is known to both have a gain-of-function germline feminizing ability and also to be subject to co-suppression when introduced ectopically into *C. elegans* (Francis et al. 1995a, Jones and Schedl 1995, Dernburg 2000), these feminized animals may also over-express the *Cbr-gld-1* transgene, thus causing a loss-of-function phenotype.

Penetrant rescue of the *C. elegans gld-1* null mutation *q*485 by wild-type *C. briggsae* GLD-1 under control of wild-type *C. briggsae* transcriptional and translational regulatory sequences strongly suggests that Cbr-GLD-1 possesses all the necessary biochemical activity of Ce-GLD-1 for proper germline development in *C. elegans*. Furthermore, it shows that upstream and downstream regulatory sequences of *Cbr-gld-1* are competent to dictate the proper temporal and spatial expression patterns of *C. elegans* GLD-1.

gld-1 acts across Caenorhabditis to control progression through oogenic meiosis

GLD-1 belongs to a widely conserved protein family (see Vernet and Artzt 1997 for review), and we wanted to determine *gld-1*'s function in other *Caenorhabditis* species. In particular, we sought to uncover a possible sex determination phenotype for *gld-1* in male/female species of *Caenorhabditis* to determine if *gld-1* might have had an ancestral role in this process, or perhaps was instead independently recruited by *C. elegans* and *C. briggsae* for germline sex determination, albeit for different major functions.

All seven *Caenorhabditis* species we investigated produced GLD-1 protein as judged by Western blot analysis. As shown in Figure 18, we find that the pattern of high GLD-1 levels in XX

animals and low levels in XO animals seen in *C. elegans* holds true across these seven species, regardless of mating system (Jones et al. 1996 and data not shown). This suggests that GLD-1 may have a role in the females of gonochoristic species, but may not be important in males.

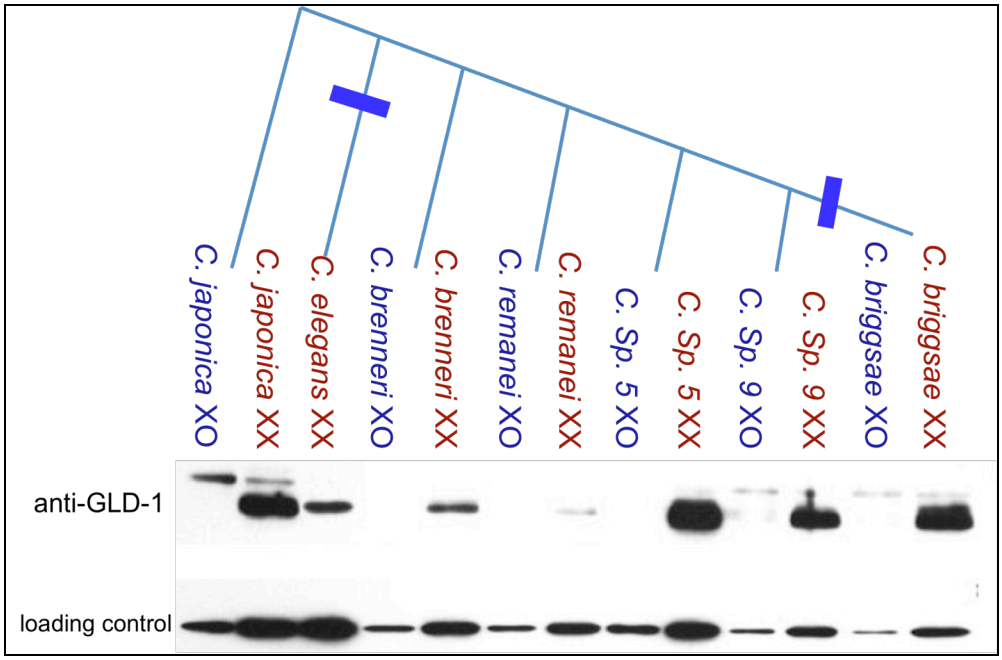


FIGURE 18. Western blot analysis of GLD-1 in females/hermaphrodites (karyotype XX) and males (karyotype XO) from 7 species of *Caenorhabditis*, including undescribed species Sp. 5 and Sp 9. anti-GLD-1 antibody is chicken 1026 anti-GLD-1, described in dissertation Chapter 2; tubulin is a "loading control." Crude protein was made from 40 adults per lane. Given above species names is a cladogram depicting the evolutionary relationships of taxa. Thick blue bars indicate a gain of hermaphroditism in *C. elegans* and *C. briggsae*.

We investigated a potential role for *gld-1* in sex determination of gonochoristic *Caenorhabditis* species by using RNAi to knock down *gld-1* expression in *C. remanei*, *C. brenneri*, *C. japonica*, and F1 hybrids from a *C.* Sp. (EG5826) female x *C. briggsae* (AF16) male mating. The gonochoristic Sp. 9 is the most closely related species to *C. briggsae* known (Marie-Anne Felix, personal communication). F1 XX hybrids from a Sp. 9-*C. briggsae* mating are true females, and yet have one copy of '"hermaphrodite" genes from *C. briggsae* (Gavin Woodruff, personal communication). We wanted to see whether in this sensitized, hybrid genetic background, a sex determination role for *gld-1* could be revealed.

In *C. briggsae*, we found that injecting low concentrations of *gld-1* RNA (~0.5ug/ml) produced weak phenotypes in the progeny of injected mothers, such as aberrant oogenesis and subsequent delayed fertility. At higher RNA concentrations (~3ug/ul) however, we were able to produce the loss-of-*gld-1* phenotypes seen in Nayak et al. 2005 (Figure 19 (E)): Mog germlines and tumors. Indeed, the injected mothers of strongly affected progeny often developed whole germline mitotic tumors themselves between 1 and 2 days post-injection.

When species-specific *gld-1* dsRNA (2-3ug/ul) was injected into mated females of four gonochoristic *Caenorhabditis* species, only two obvious mutant phenotypes were recovered: germline tumors and morphologically aberrant oocytes. We singled F1 female animals away from their brothers at the L4 stage for each species to check for self-fertility, but failed to recover any self-fertile animals. Additionally, DIC examination of 50 affected females per species did not reveal any sperm (Figure 19 (A-D)). Examination of 50 males from plates with (at least nearly) 100% affected female siblings only revealed normal-looking male germlines, though we did notice *C. japonica* males with expanded regions of primary spermatocytes (data not shown). We also note that *gld-1* RNAi mutant phenotypes in *C. brenneri* were of very low penetrance compared to other species injected, as only about 10% of progeny were affected per injected mother (even if that mother herself developed a germline tumor).

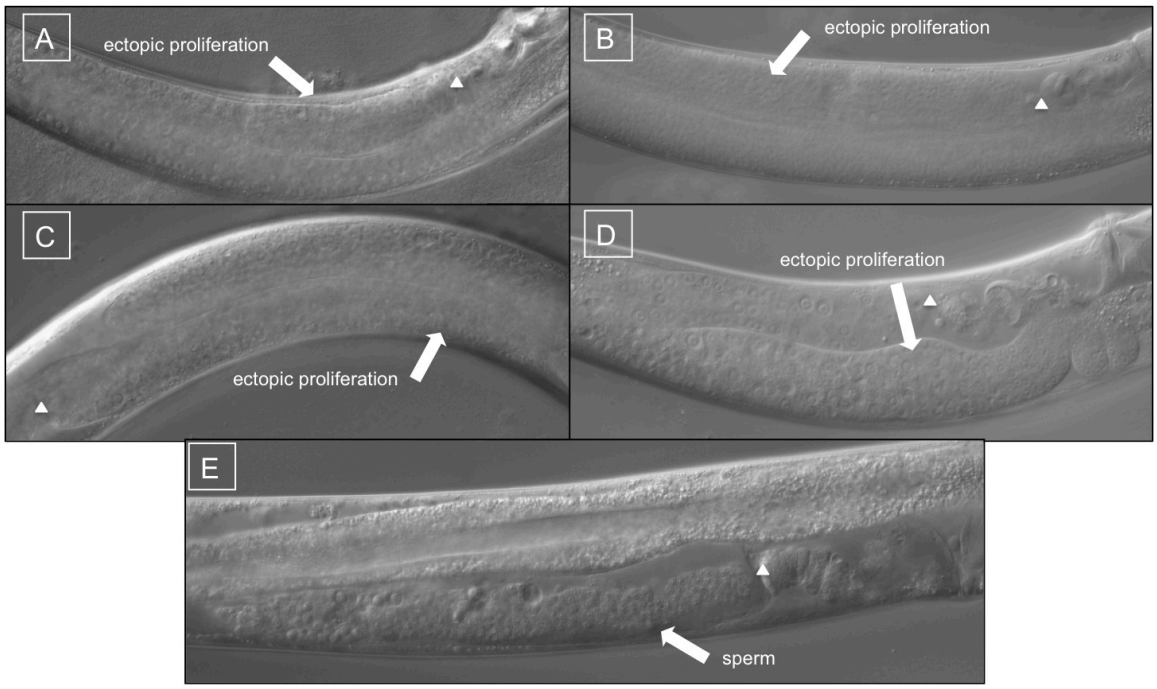


FIGURE 19. DIC micrographs of XX-karyotype F1 progeny from *gld-1* RNAi injected mothers of (A) Sp. 9–*C. briggsae* hybrid; (B) *C. remanei*; (C) *C. brenneri* ; (D) *C. japonica*; (E) *C. briggsae*. All germlines except *C. briggsae* develop a germline tumor and no sperm, whereas *C. briggsae* affected germlines contain only sperm. White triangles indicate the proximal end of the gonad, near the vulva.

Given *gld-1*'s role in oogenesis progression in both *C. elegans* and *C. briggsae* (Francis et al. 1995a, Nayak et al. 2005, this work), it was possible that the germline tumors formed in female *C. remanei, C. brenneri, C. japonica,* and *C.* sp. 9/*C. briggsae* F1 hybrids with *gld-1* RNAi were made of oogenic cells that had returned to the mitotic cell cycle. In order to confirm this, we made protein from affected worms of each species and performed a Western blot analysis with antibodies to the oocyte yolk receptor RME-2 and a sperm-specific protein, MSP. We found that all wild-type females and RNAi-affected animals from these different species possess only RME-2 and not MSP (data not shown), whereas wild-type *C. briggsae* hermaphrodites and *Cbr-gld-1(nm68)* mutants only show a signal for MSP. These results demonstrate that *gld*-1 knockdown in females of gonochoristic species results in creation of oocyte-fated tumors. Thus

gld-1's role in oocyte progression is widely conserved in *Caenorhabditis*. However, we find that only *C. briggsae* suffers from a sexual transformation of gamete fate, i.e., of oocytes into sperm.

DISCUSSION

nm41, nm64, and *nm68* are alleles of *C. briggsae gld-1*

Through both forward and reverse genetic screening we obtained 3 alleles of *C. briggsae gld-1*: *nm41* is a nonsense mutation in amino acid 14, *nm64* is a nonsense mutation within the evolutionary conserved region downstream of *gld-1*'s KH domain, and *nm68* is a deletion that removes 55% of the *gld-1* coding region, including the KH domain and conserved downstream region. By genetic linkage, complementation tests, and DNA sequencing, we conclude that mutations in *gld-1* are responsible for the mutant phenotypes in *nm41*, *nm64*, and *nm68*.

We might expect that such an early stop codon as *nm41* would be a null mutation. However, there is a methionine codon 70 positions C-terminal to the *nm41* sequence lesion that is conserved between *C. briggsae* and *C. elegans*; in Jones et al. 1996, the authors speculate that the corresponding *Ce-gld-1* mutation *q93oz50* is a weak loss of function perhaps because that downstream transcription start site is utilized. The deletion allele *nm68* preserves almost all of the evolutionarily conserved region upstream of the KH domain (the "Quaking 1 domain"), but removes the KH domain itself and also the downstream conserved region (the "Quaking 2 domain) (Ebersole et al. 1996, Vernet and Artzt 1997). As amino acids surrounding the KH domain are important for both GLD-1 homodimerization and RNA binding, it is likely that *nm68* is a null mutation (Chen et al. 1997, Lin et al. 1997, Lehmann-Blount KA et al. 2005, Beuck et al. 2010). RT-PCR experiments show that *nm41*, *nm64*, and *nm68* mutants all produce mRNA products, but none of the alleles produce detectable GLD-1 protein by Western blot, though we

assume that because their phenotypes can be distinguished from one another and because analogous mutations in *C. elegans gld-1* are known to have different molecular properties, these *C. briggsae* alleles are not all protein nulls.

gld-1 mutants cause masculinization of the hermaphrodite germline in C. briggsae but feminization in C. elegans

Mutations in *C. briggsae gld-1* cause a range of related germline mutant phenotypes including masculinization of the hermaphrodite germline, proximal germline tumors, full germline tumors that begin beyond the wild-type meiotic entry point, and combinations of these mutant phenotypes. Additionally, as primary spermatocytes are often observed in *gld-1* mutant gonads a few days after reaching adulthood, but never in wild-type animals at this point, we posit that there is also a meiotic delay in *C. briggsae gld-1* mutants. This may cause or contribute to their proximal tumor formation, perhaps by allowing inappropriate interactions between cells that are delayed in meiotic entry and the sheath cells of the somatic gonad (Killian and Hubbard 2004, Killian and Hubbard 2005). Interestingly, proximal tumors are not described for *C. elegans gld-1* mutants.

There are both important differences and similarities between the mutations in *C. briggsae gld-1* described here and the *C. elegans gld-1* mutations described in Francis et al. (1995a, 1996b) and Jones et al. (1996). Strikingly, though *gld-1* mutants of both species can form germline tumors, the sexual fate of these mutant germlines is opposite in the two species. In *C. elegans*, Francis et al. showed with extensive double mutant analysis that *C. elegans gld-1* tumors only form when the germline is set to produce oocytes. Indeed, these germline tumors stain strongly with an antibody to the oocyte yolk receptor RME-2 and not with an antibody to the Major Sperm Protein (Nayak et al. 2005, this work). Furthermore, *gld-1* is also

haploinsufficient for specification of the sperm fate in *C. elegans* (Francis et al. 1995a). Thus, *Ce-gld-1* is necessary to specify hermaphrodite sperm in *C. elegans*.

In *C. briggsae gld-1* mutants, however, we find that except for weak RME-2 staining in *nm64* mutants (which may be the result of a shared gain-of-function property with *C. elegans* (discussed below)), *gld-1* germlines stain only with an antibody to MSP and not RME-2. Considering the overproduction of sperm in *Cbr-gld-1* mutant germlines and their lack of staining with the anti-yolk-receptor antibody, we conclude that the major sex determination role of *C. briggsae gld-1* is to specify the oocyte fate in hermaphrodites. This is opposite to *gld-1*'s role in C*. elegans*, and it coincides with what Nayak et al. (2005) concluded from *Cbr-gld-1* RNAi experiments. This species difference in sexual phenotype is likely a product of the independent involvement of *gld-1* in the evolution of hermaphrodite sex determination in *C. elegans* and *C. briggsae*. As *gld-1* in both species has complicated genetics, with multiple developmental functions and many protein binding partners and messenger RNA targets (Clifford et al. 2000, Lee and Schedl 2001, this work, dissertation Chapter 2), *gld-1* may also have other effects on sex determination that are similar to, or may contrast with, our inferences here (see for instance Kim et al. 2009). Nevertheless, there is a single, overt germline sex determination phenotype that results from the *Ce-gld-1* null mutation and the putative null *Cbr-gld-1* mutation: feminization of the germline in *C. elegans* mutants and masculinization of the germline in *C. briggsae* mutants.

A lesser, but interesting, difference in *gld-1* mutant phenotypes between *C. elegans* and *C. briggsae* is the range of variability exhibited by mutants of the two species. In general, mutations in *C. elegans gld-1* possess a characteristic, penetrant phenotype such that the first 31 mutants described were able to be grouped into six phenotypic classes (Francis et al. 1995a). However, each of the three *Cbr-gld-1* mutant alleles described here has a range of phenotypes,

and that range is the same for all alleles, differing only in phenotype frequency among them; there is no “tumorous class” or “pachytene arrest class” class as there is in *C. elegans*. Perhaps with additional *C. briggsae* mutant alleles, we would come to find a pattern like in *C. elegans*. More interestingly, however, these differences in phenotypic variation may also be evidence of independent recruitment into gametogenesis, discussed below.

C. briggsae GLD-1 can act as a translational repressor as in C. elegans

The high sequence identity between *C. elegans* and *C. briggsae gld-1*, as well as their common roles in directing oocyte progression and repressing the translation of RME-2 as shown by Nayak et al., suggests that GLD-1 in both species has similar biochemical activities and patterns of action. Why , then, is the major sex determination role of *gld-1* in *C. elegans* to promote hermaphrodite spermatogenesis but to promote hermaphrodite oogenesis in *C. briggsae*? One hypothesis is that differences in the messenger RNA targets and/or protein binding partners of *C. elegans* and *C. briggsae* GLD-1 might be responsible for their different major sex determination functions, as opposed to changes in GLD-1 protein function itself.

To test this idea, we demonstrated rescue of the *C. elegans gld-1* null allele *q485* with the *C. briggsae gld-1* locus, including 1000 base pairs of upstream and downstream regulatory sequence. As *gld-1* has multiple roles in germline development in both *C. briggsae* and *C. elegans*, it is notable that *C. briggsae gld-1* can rescue all of Ce-GLD-1’s obvious functions, including specification of hermaphrodite spermatogenesis, oogenic progression through meiosis, and proper oocyte formation. Perhaps even more surprising is that despite more that 40 million years of independent evolution between them, the regulatory sequences of *Cbr-gld-1* are competent to drive the proper temporal and spatial expression and dictate the proper amount of Ce-GLD-1 for rescue. Thus we conclude that Cbr-GLD-1 possesses the same

biochemical activity as Ce-GLD-1, though this cross-species rescue does not show that Cbr-GLD-1 must be acting identically within *C. briggsae* as it does in *C. elegans*, only that it is capable of carrying out the same functions as *C. elegans* GLD-1. Nevertheless, this result supports the idea that the difference in major sex determination phenotype between *C. elegans* and *C. briggsae gld-1* mutants is due to differential messenger RNA targets and/or protein binding partners of GLD-1, as opposed to changes in the GLD-1 protein itself or its regulation. A investigation of this possibility is the theme of dissertation Chapter 2.

Cbr-gld-1 tumors may result from a hermaphrodite-specific spermatogenesis defect

Cbr-gld-1 mutants have germlines that may be fully masculinized, fully tumorous, or have some mixture of sperm plus tumorous cells. Examination of mutant germlines suggests that just as in *C. elegans*, *Cbr-gld-1* tumors result from cells that have entered meiosis returning to the mitotic cell cycle. This is indicated by formation of tumors only once cells have entered mitosis and by the presence of mitotic chromosomal figures mixed with pachytene-like cells and primary spermatocytes adjacent to them.

The *C. elegans gld-1* mutant germline tumors strongly express the oocyte marker RME-2. This is consistent with *Ce-gld-1* being necessary for both specification of hermaphrodite spermatogenesis and also progression through oogenic meiosis, such that cells in homozygous mutant germlines develop the oocyte fate, yet can't complete oogenesis and instead form a germline tumor. In contrast, tumorous cells of *C. briggsae gld-1* mutant germlines do not stain with either the RME-2 or MSP markers of sexual fate. This demonstrates that XX germ cells in *C. briggsae gld-1* mutants are not feminized, or are at least not feminized in the same way as in *C. elegans*. It leaves open the possibility, though, while that some cells of *Cbr-gld-1* mutant

germlines are committed to the male fate (i.e., sperm), others remain uncommitted or have an undetectable oocyte character.

While considering the sexual fate of cells that stain neither with anti-RME-2 nor with anti-MSP antibodies, we should distinguish between the fate of cells as they first suffer the loss of *gld-1* versus the "fate" of mitotically dividing, tumorous cells that have already exited meiosis in *Cbr-gld-1* mutant germlines. Given the general masculinized nature of *C. briggsae gld-1* mutant gonads, three explanations seem likely with regard to the sexual fate of the aberrant mitotic cells: 1) they undergo meiosis fated to become sperm, and still retain some sperm character upon re-entry into mitosis (analogous to the RME-2-expressing tumor cells in *C. elegans*), but do not express MSP in particular; 2) they undergo meiosis fated to become sperm, but re-enter mitosis and subsequently loose sexual identity (as hypothesized for *puf-8*(lf) spermatocytes in *C. elegans* (Subramanian and Seydoux 2003); or 3) they do not have a sexual fate to begin with and thus are sexually uncommitted upon their return to mitosis. It is also formally possible that the masculinizing sex determination force from *Cbr-gld-1*(lf) is too weak in mutant germlines to specify that all cells develop as sperm such that the remaining cells actually come to have oocyte character. If this were true, then cells not masculinized by loss of *Cbr-gld-1*, but that do not have enough feminizing activity to express RME-2, may not be able to proceed through meiosis I and either become arrested or return to mitosis.

Supposing a potent role for *Cbr-gld-1* in sex determination, we hypothesize that loss of *Cbr-gld-1* in the hermaphrodite germline sets the sexual fate of all gametes to sperm. Then, *Cbr-gld-1* is also required for reliable progression through hermaphrodite (but not male) spermatogenesis, such that as sperm-fated cells are undergoing spermatogenesis in *Cbr-gld-1* mutant germlines, they often complete spermatogenesis, but other times cannot; some cells of the latter suffer meiotic arrest or delayed spermatogenetic progression, and others return to

the mitotic cell cycle, either truly dedifferentiated or simply not expressing MSP. This hypothesis is consistent with the *Cbr-gld-1* RNAi phenotype, in which masculinized mutant animals possess a germline full of mature sperm, but very rarely manifest the meiotic mutant and tumor phenotypes seen in true *Cbr-gld-1* mutants (Nayak et al. 2005, this work); as RNAi likely produces a weaker loss of function phenotype than the *Cbr-gld-1*(nm68) deletion mutant, we can perhaps infer that losing a little *gld-1* activity causes germline masculinization, whereas a more severe loss of function affects spermatogenic meiosis. This proposed pattern of activity is the same as for *C. elegans gld-1* in oogenic meiosis: weak loss of function mutations in *Ce-gld-1* produce small, abnormal oocytes in hermaphrodites, whereas stronger loss of function mutations cause oocyte-fated cells to abort meiosis and return to the mitotic cell cycle.

A *Cbr-gld-1* role in spermatogenic meiosis would stand in contrast to *C. elegans gld-1*. In *C. elegans*, double mutant combinations of *Ce-gld-1*(lf) with germline masculinizing mutations (like *mog-1(lf)*; *gld-1(null)* and *tra-3*(lf-germline-specific); *gld-1(null)*) have normal spermatogenesis. Additionally, as described above, both *C. elegans gld-1*(null) and *C. briggsae gld-1 nm41* and *nm64* XO wild-type animals make sperm normally. Thus, *C. elegans gld-1* is not needed for proper hermaphrodite spermatogenesis, nor is *gld-1* needed for meiosis in the males of either species. Therefore, a potential role for *gld-1* in spermatogenesis would be a *C. briggsae* hermaphrodite-specific feature.

We tried to demonstrate this hypothesized role for *gld-1* in spermatogenesis in *C. briggsae* hermaphrodites by constructing double mutants with *Cbr-gld-1* mutations and the strong loss of function allele *Cbr-tra-2*(nm1), which masculinizes the soma and germline of XX animals. However, double mutants possessed *tra-2*-like germlines with normal spermatogenesis. From these results, we can infer that either *Cbr-gld-1*(lf) germline tumors are not developed from spermatogenic cells, or that perhaps *tra-2* pseudomales are masculinized enough to have a

gld-1 mutant response like XO males –that is, normal spermatogenesis. We think this latter explanation is possible, because though *C. elegans gld-1*(lf) mutations cause some feminization in XX Mog mutants (i.e., mutations like *Ce-fem3*(gf) and *Ce-mog-1*(lf) that alone cause only masculinization of the hermaphrodite germline, not the soma), 100% of the germlines of *gld-1*(lf); *tra-2*(lf) and *gld-1*(lf); *tra-3*(lf) double mutant pseudomales make only sperm (Francis et al. 1995b). These data are consistent with *gld-1* having no effect when worms are highly masculinized.

The potential involvement of *Cbr-gld-1* in spermatogenic meiosis might be a consequence of the independent acquisition of hermaphroditism in *C. briggsae* and *C. elegans*. Given our knowledge of the *Caenorhabditis*-wide role of *gld-1* in oogenic meiosis (this work), it seems that *gld-1* is always expressed in the early meiotic cells of XX animals, likely serving as a conserved early translational repressor of oocyte/maternal RNAs across species (Lee and Schedl 2001, dissertation Chapter 2). This would likely have been true in the gonochoristic ancestors of *C. briggsae* and *C. elegans* as well. Later, as these two species independently evolved the ability to form sperm in a female body, they may or may not have eliminated (unnecessary) *gld-1* activity from their newly acquired meiotic ability: spermatogenesis. Interestingly, both the relative delay of spermatogenesis in *C. briggsae* hermaphrodites compared to *C. elegans* and also the relatively simultaneous formation of both types of gametes in *C. briggsae* perhaps makes it more likely that *Cbr-gld-1* could have an effect in both spermatogenesis and oogenesis in *C. briggsae*. The similar return-to-meiosis defects observed in both single mutant *Cbr-gld-1* germlines and in double RNAi mutants when *Cbr-gld-1* and the sperm-promoting gene *fog-3* are knocked down together are consistent with this idea (Nayak et al. 2005, this work).

The GLD-1 expression pattern in *C. elegans* hermaphrodites is consistent with its involvement in the progression of only oocyte-fated cells through meiosis, not sperm: in

presumptive sperm cells of the *C. elegans* hermaphrodite larval germline, GLD-1 is only detected in early meiotic prophase I; by the time these cells enter pachytene, they no longer have GLD-1 staining, consistent with a limited role of *Ce-gld-1* in spermatogenic meiosis (Jones et al. 1996). If *C. briggsae* GLD-1 is necessary for proper hermaphrodite spermatogenesis, then in contrast to *C. elegans*, we would expect to detect GLD-1 in pachytene cells of presumptive hermaphrodite sperm in *C. briggsae*. We are currently testing this possibility.

Oocytes, not oocyte tumors, are found in Cbr-gld-1; tra-1(nm2) mutant germlines

To place *Cbr-gld-1* in the *C. briggsae* germline sex determination pathway, we performed epistasis analysis between different *Cbr-gld-1* mutations and the loss-of-function allele *Cbr-tra-1*(nm2), which masculinizes the hermaphrodite soma and allows both abnormal amounts of sperm to be made and then oocyte production in the germline. However, we were surprised to find recognizable, good quality oocytes in the germlines of *tra-1* mutants in combination with either *Cbr-gld-1 nm68* or *nm41* (where one-fourth of Tra animals scored were also homozygous mutant for *gld-1*(lf)), because Nayak et al. has shown that *Cbr-gld-1* is necessary for keeping oocyte-fated cells committed to meiosis. The good-quality oocytes we see in *Cbr-gld-1*(*nm41* or *nm68*); *tra-1(nm2)* germ line also is contrary to the comparable double mutant phenotype of *C. elegans gld-1*(*q485* or *q268*)*; tra-1*(*e1834*) homozygotes, where mutant animals contain both sperm and ectopically proliferating cells (Francis et al. 1995b).

That we see oocytes, and not oocyte tumors, in *Cbr-gld-1*(nm68 or nm41); *tra-1*(nm2) animals may indicate that in *C. briggsae,* tumors with oocyte character observed in the *gld-1; fog-3* double RNAi mutants of Nayak et al. (2005) are perhaps dependent upon *tra-1* itself. Alternatively, the differences in *gld*-1(lf); *tra-1*(lf) double mutant phenotypes in the two species could be a consequence of differences in the *tra-1* alleles in question: Francis et al.'s (1995b) *Ce-*

tra-1(e1834) is a deletion allele that removes most of the Ce-*tra-1* coding region, whereas *Cbr-tra-1*(nm2) is a nonsense mutation 50% through the coding region that preserves *Cbr-tra-1*'s zinc finger domain. These two alleles may produce *tra-1* mutant oocytes with different character. A third explanation for the oocytes found in *Cbr-gld-1;tra-1*(lf) animals is that in *C. briggsae*, tumor formation may not be possible in XX pseudomales. This restriction on the ability of *Cbr-gld-1* mutations to cause germline tumors would be at odds with *C. elegans gld-1*, which produces oocyte tumors in feminized germlines regardless of somatic sex (Francis et al. 1995b); it may also be consistent with the lack of tumor formation in *Cbr-gld-1*(lf); *tra-2*(lf) XX pseudomales.

gld-1's role in progression through oogenesis is conserved in Caenorhabditis

It seemed plausible that *gld-1* might have already served to repress XX sperm production in the gonochoristic ancestors of *C. briggsae*. If so, one or more extant gonochorist might retain this. If not, then perhaps *gld-1* was independently recruited by *C. elegans* and/or *C. briggsae* for hermaphrodite sex determination. All known members of the Elegans group of *Caenorhabditis* (Kiontke and Sudhaus 2006) produce GLD-1 protein, and we find that the pattern of strong expression in XX animals and weak in XO males observed in *C. elegans* by Jones et al. 1996 is conserved across a group of 7 *Caenorhabditis* species.

Next, we used RNAi to knock down *gld-1* in *C. remanei*, *C. brenneri*, *C. japonica* and in F1 XX female hybrids of a mating between the gonochorist Sp. 9 and *C. briggsae*. We recovered only two obvious *gld-1* RNAi phenotypes from this work: germline tumors and morphologically aberrant oocytes in XX animals. By Western blot analysis, we determined that these germline tumors express the oocyte yolk receptor protein RME-2, and we detected no sperm by DIC microscopy or by Western blot. Conversely, XO male germlines of all species appeared normal, and the brothers of affected XX siblings were able to sire progeny with unaffected conspecifics.

Together with work in *C. elegans* (Francis et al. 1995b) and *C. briggsae* (Nayak et al. 2005), these results show that directing meiotic progression through oogenesis is a conserved property of *gld-1* in *Caenorhabditis.*

However, we cannot reveal a sex determination function for *gld-1*, i.e. transformation of oocytes into sperm, in any species other than *C. briggsae* by this assay. This is true even in Sp.9-*C. briggsae* hybrids, which have half a genome of '"hermaphrodite" genes from *C. briggsae*. These observations are consistent with two hypotheses: first, that *gld-1* has no role in sex determination in any species other than *C. briggsae* or *C. elegans*; second, that *gld-1* has a sex determination role to specify the oocyte fate in gonochoristic species, but females of these species may be canalized for oocyte production such that losing *gld-1* through RNAi by itself has no demonstrable effect.

We do note one piece of evidence of an ancestral sex determination role of *gld-1* in *Caenorhabditis*: the potentially shared gain-of-function feminizing character of both *Ce-gld-1 (q286* and *q395* alleles) and *Cbr-gld-1(nm64)* mutations. As described, these alleles share the same nonsense mutation, and though the major sex determination mutant phenotypes are opposite in the two alleles, *nm64* does produce low levels of RME-2 as judged by Western blot and antibody staining of dissected gonads. Francis et al. (1995a) showed that while loss-of-function for oogenesis, the feminizing activity of *q286* is a gain-of-function character. If a shared gain-of-function property can explain the weak RME-2 staining we see in *nm64*, this might suggest that *gld-1* was involved somehow in sex determination before *C. elegans* and *C. briggsae* ancestors speciated.

CHAPTER 2:

Characterizing Cbr-GLD-1 at different biological scales: its role in sex determination, genome-wide identification of mRNA targets, and evolutionary position within STAR proteins

ABSTRACT

STAR-domain proteins are widely conserved regulators of RNA. In *C. elegans*, the STAR protein GLD-1 is a germline-specific translational repressor with pleiotropic effects on germline development, including the specification of hermaphrodite sperm in an otherwise female body. Within *Caenorhabditis*, *C. briggsae* also produces hermaphrodites; however, *C. elegans* and *C. briggsae* independently evolved hermaphroditism from different male/female ancestors. *C. briggsae* possess a GLD-1 ortholog, and like Ce-GLD-1, it functions in multiple germline events, including sex determination. However, we previously confirmed that *gld-1* has an opposite major sex determination role in *C. elegans* and *C. briggsae*, and showed further that Cbr-GLD-1 can rescue the *Ce-gld-1*(null) mutant phenotype. This suggests that a change(s) in the protein binding partners and/or messenger RNA targets of GLD-1 is responsible for its different sex determination roles in *C. briggsae* and *C. elegans*, rather than a change in GLD-1 function or regulation itself. To identify the messenger RNA targets of Cbr-GLD-1 that might be responsible for sex determination in *C. briggsae* hermaphrodites, we used an *in vivo* genome-wide approach to simultaneously identify mRNAs associated with Cbr-GLD-1, including potential sex determination targets. We identify 802 putative Cbr-GLD-1 mRNA targets and confirm the specificity of this gene set with qRT-PCR and RNAi. To understand the multifaceted roles of GLD-1 in its evolutionary context, we created a phylogeny of STAR proteins across metazoans. We find that GLD-1 lies within a nematode-specific expansion of the STAR family, and that GLD-1 is part of a larger 'super-clade' containing vertebrate Quaking and Drosophila How/Who.

INTRODUCTION

STAR proteins (for **s**ignal **t**ransduction and **a**ctivation of **R**NA metabolism) are widely conserved regulators of RNA. Research on family members in Drosophila, mice, Xenopus, and *C. elegans* has revealed biological roles for STAR proteins in cell division, gametogenesis, apoptosis, and embryonic and larval development.

In *C. elegans*, the STAR protein GLD-1 is a germline-specific, pleiotropic translational repressor of mRNAs (for instance, Francis et al. 1995a and 1995b, Lee and Schedl 2001, Marin and Evans 2003, Mootz et al. 2004). Ce-GLD-1 is involved in multiple events necessary for proper nematode germline development, including the mitosis/meiosis decision of germline stem cells, meiotic progression of oocyte-fated cells, and specification of *C. elegans* hermaphrodite sperm in an otherwise female body. *C. elegans* GLD-1 controls the production of hermaphrodite sperm for germline sex determination by binding to specific elements in the female-promoting *tra-2* mRNA 3'UTR, causing its translational repression by an unknown mechanism (Goodwin et al. 1993, Jan et al. 1999, Ryder et al. 2004).

Within the Elegans group of *Caenorhabditis* nematodes, there is another species that also produces hermaphrodites, *C. briggsae*. Surprisingly, phylogenetics has revealed, and genetic evidence supports, that *C. elegans* and *C. briggsae* each evolved the ability to produce hermaphrodites independently from different male/female ancestral species (Kiontke et al. 2004, Nayak et al. 2005, Hill et al. 2006, Guo et al. 2009). *C. briggsae* possess a GLD-1 ortholog, and like Ce-GLD-1, it is expressed in the germline; can translationally repress at least one target in common with Ce-GLD-1, *rme-2*; and is involved in multiple germline events, including meiotic progression of germ cells and sex determination. However, we previously confirmed with genetic mutations in *Cbr-gld-1* that *gld-1* has an opposite major sex determination role in *C.*

elegans and *C. briggsae*: loss of *Ce-gld-1* results in germline feminization, whereas loss of *Cbr-gld-1* results in germline masculinization (Jones et al. 1996, Nayak et al. 2005, dissertation Chapter 1). In a cross-species rescue experiment, we also previously demonstrated that the *Cbr-gld-1* coding sequence plus regulatory sequences can rescue all visible aspects of the *Ce-gld-1*(null) mutant phenotype, including progression through meiosis, proper oocyte formation, and specification of hermaphrodite spermatogenesis. This suggests that a change(s) in the protein binding partners and/or messenger RNA targets of GLD-1 is responsible for its different sex determination roles in *C. briggsae* and *C elegans* rather than a change in GLD-1 function or regulation itself.

To identify the messenger RNA targets of Cbr-GLD-1 that might be responsible for its sperm-promoting role in *C. briggsae* hermaphrodites, we used an *in vivo* genome-wide approach, RIP-chip ("RNA immunoprecipitation microarray chip"), to simultaneously determine the many mRNAs associated with Cbr-GLD-1, including its potential sex determination targets. Immunoprecipitating endogenous Cbr-GLD-1, we identify 802 genes (with a false discovery rate <2%) that are enriched in anti-GLD-1 immunoprecipitations compared to mRNAs recovered from both mock IPs and total mRNA. Western blot analysis, quantitative RT-PCR, and a large RNAi screen of enriched mRNAs confirms the identity of these mRNAs as putative Cbr-GLD-1 targets.

To understand the multifaceted roles of *C. elegans* and *C. briggsae* GLD-1 in their evolutionary context, we created a phylogeny of STAR proteins across metazoans. We find that well-studied STAR proteins, like Quaking and SAM68, form distinct, well-supported clades, and that GLD-1 belongs to a 'super-clade' containing Quaking, Drosophila How/Who, and *C. elegans* ASD-2. Further, we hypothesize that ancestral proteins belonging to two super-clades of STAR proteins, the Quaking/How/GLD-1 clade and SAM68-like clade, were present before the split of cnidarians and bilaterians at the base of the animal tree.

METHODS

anti-GLD-1 antibody production

In consultation with Open Biosystems, two potentially antigenic regions of *C. briggsae* GLD-1, each ~20 amino acids long, were identified that were also common to GLD-1 orthologs in *C. elegans* and *C. remanei* but not other homologs. We used chicken for antibody production to take advantage of the large supply of IgY found in chicken eggs and the fact that with boosting, antibody can be purified through eggs for the life of the chicken. (We also thought an antibody developed in a less common animal like chicken would maximize its utility in immunohistochemistry experiments involving multiple proteins (e.g., double antibody stainings)). Two chickens were injected with a synthetic peptide for each antigenic region. We obtained one affinity-purified polyclonal chicken antibody specific to GLD-1 (verified in dissertation Chapter 1 and in the *Results* below), designed to residues 16-36 of Cbr-GLD-1.

Immunoprecipitation of GLD-1 *and recovery of RNA*

To isolate messenger RNA targets of Cbr-GLD-1*,* we immuoprecipitated GLD-1 from *C. briggsae* young adults using the chicken anti-GLD-1 antibody described above. We used a modified version of an immunoprecipitation/mRNA recovery protocol kindly provided by Aaron Kershner (Kimble laboratory, University of Wisconsin; Kershner and Kimble 2010), adjusted for our chicken antibody and incorporating other ideas from the literature described below.

Synchronous, largely hermaphroditic populations of AF16 were grown on 15cm NGM plates at 20°C seeded thickly with OP50. 2-4 separate populations of worms were grown in parallel, each kept separate through growth and processing to comprise biological replicates. When their average age was young gravid adult, animals were rinsed off their plates and washed repeatedly in M9 (for more than 30 total minutes in order to clear bacteria from the

worm gut). We recovered ~1ml of worm "pellet" (by gravity settling or light spinning) per replicate, and each pellet was washed once in buffer A (20mM Tris pH 8.0, 150mM NaCl, 1mM EDTA), twice in lysis buffer (20mM Tris pH 8.0, 150mM NaCl, 1mM EDTA, 0.1% NP-40, 0.2mg/ml heparin, plus 1x EDTA-free Mini Complete Protease Inhibitor Cocktail (Roche), 2mM DTT, 200 U/ml recombinant RNasin (Promega)), and then finally resuspended 1:1 in lysis buffer.

Next, worms were poured into a cold mortar and ground in liquid N2 into a fine powder. Worm power was thawed on ice, passed ~15 times through a dounce homogenizer on ice, and then brought up to 2ml with lysis buffer. Worm mixtures were spun at 10,000 g for 10 minutes at 4º C to pellet insoluble debris, and supernatants were pre-cleared with goat anti-IgY agarose-coupled beads ("PrecipHen," Aves Labs, Tigard, Oregon) by rotation for 30 minutes at 4°C. Finally, half the pre-cleared lysate of each biological replicate was added to 15ug chicken anti-GLD-1-bound PrecipHen and half to 15ug total IgY-bound PrecipHen (using unconjugated total IgY from Jackson Immunoresearch, #003-000-003), and rotated for ~8 hours at 4°C.

To harvest the RNA bound to GLD-1, beads were spun down after immunoprecipitation and washed 10 minutes tumbling at 4°C in lysis buffer and then 4x10 minutes tumbling at 4°C in wash buffer (20mM Tris pH 8.0, 150mM NaCl, 1mM EDTA, 10% glycerol, 0.1% NP-40, 1mM DTT, 10 U/ml RNasin). We released the RNA from GLD-1 by twice adding 500ul TRI Reagent (Ambion) to each tube and rocking/shaking for 30 minutes. We then added 200ul chloroform to the TRI Reagent-supernatant and performed a modified phenol:chloroform extraction and ethanol precipitation using Qiagen RNeasy columns. RNA was eluted in RNase-free water.

During the experiment, samples were removed after pre-clearing and after the immunoprecipitation to check for both the quality of RNA recovered by agarose gel and the effectiveness of the immunoprecipitations by Western blot.

Microarray and experimental design

We obtained an oligo probe set for all predicted *C. briggsae* protein genes (WormBase *C. briggsae* version cb25.agp8) as a very kind gift from Dr. Itai Yanai (Technion-Israel Institute of Technology; Yanai and Hunter 2009). In this probe set, probes are 3'-biased, and 100% of genes are represented by at least 1 probe, 98% represented by at least 2 probes, and 2.6% represented by 3 probes. We added third probes where needed for the following categories of genes: candidate Cbr-GLD-1 target sex determination genes (i.e. genes with mutant germline feminizing phenotypes in *C. briggsae* and/or *C. elegans*); most other *C. elegans* and *C. briggsae* germline sex determination genes identified from the literature; all *Cbr-puf* genes (Lamont et al. 2004); other classes of germline genes we deemed 'interesting' (e.g., germline RNA binding proteins (Lee and Schedl 2006) and genes involved in RNAi/microRNA processing); 5 'positive control' genes: known targets of Cbr-GLD-1 and orthologs of known Ce-GLD-1 targets, *oma-1/2, rme-2, glp-1, mes-3, pal-1*; and 10 "negative control genes," randomly identified genes with somatic-specific or somatic–enriched expression as determined from WormBase and/or the *C. elegans* literature.

We used two Agilent 4 X 44K microarray chips for a total of 8 two-color arrays: 5 arrays for anti-GLD-1 vs. anti-IgY immunoprecipitation expression comparisons and 3 arrays for anti-GLD-1 immunoprecipitation vs. total input mRNA comparisons. We utilized the Microarray Core Facility at Washington University in St. Louis to amplify and label recovered RNA with the Kreatech aRNA Labeling Kit. The Core Facility hybridized amplified material to each array according to the scheme in Figure 1 and scanned each array for raw pixel intensity measurements.

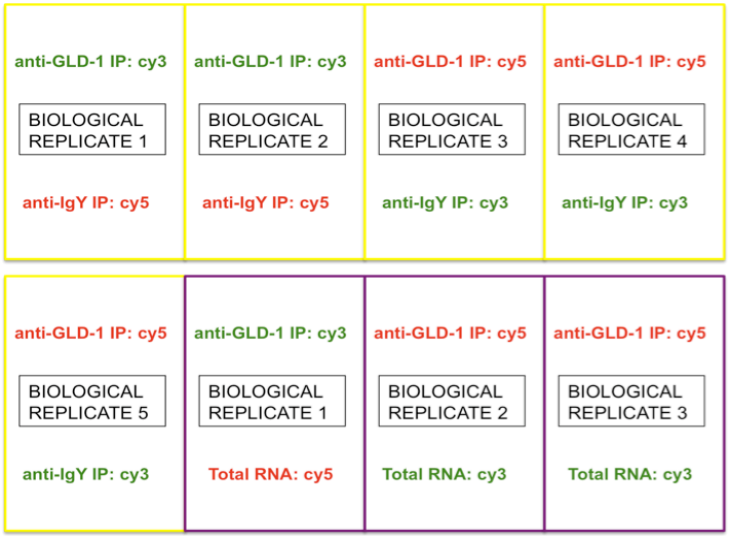


FIGURE 1. Schematic of microarray design for two different experiments, anti-GLD-1 IP vs. anti-IgY IP mRNA expression comparison and anti-GLD-1 IP vs. total input mRNA expression comparison, on two Agilent 4 x 44K arrays. Dye swaps and biological replicates are incorporated.

Analysis of microarray data

We obtained raw pixel intensity measurements from the Microarray Core Facility at Washington University in St. Louis from two expression comparisons: anti-GLD-1 IP mRNA vs. anti-IgY IP mRNA and anti-GLD-1 IP mRNA vs. total input mRNA. For data quality control, we first inspected diagnostic plots created in *limma* of foreground and background intensities vs. spatial array coordinates and MA plots of uncorrected data for each array (Wettenhall and Smyth 2004). We also made normal probability plots and histograms of raw and log2 transformed data from each experiment to assess normality and trends. We did not filter out or differentially weight any intensity values.

RIP-chip microarray data are non-standard in that even transformed and processed intensity measurements have biologically relevant skewed distributions with intensity-dependent bias. We elaborate on this in the *Results* and provide pre-processing details there.

We detected differential gene expression using two different methodologies, SAM (Tusher et al. 2001) and EDGE (Leek et al. 2006, Storey et al. 2007). In SAM, we selected 200

permutations for our two expression comparisons. The calculated s0 for the anti-GLD-1 IP vs. anti-IgY IP comparison was 0.014 (0^{th} percentile) and was 0.005 (0^{th} percentile) for the anti-GLD-1 IP vs. total mRNA comparison, meaning that the SAM "d-statistic" was essentially reduced to a t-statistic in the calculation of expression differences. In EDGE, we selected 200 iterations for both expression comparisons.

Western Blotting

Protein sample buffer/5% BME was added 1:1 to protein samples, heated at 95°C for 5 minutes, and Western blotting was performed according to standard procedures. We used a rabbit anti-GLD-1 primary antibody (Jones et al. 1996; gift of Dr. Tim Schedl, Washington University in St. Louis Medical School) at a 1:2000 dilution overnight at 4°C in 5% milk/PBST (0.1% Tween), and light-chain-specific HRP-conjugated anti-rabbit secondary antibody 1:1000 for 1-2 hours at room temperature. For loading control, we used anti-tubulin primary antibody (Sigma T9026) at 1:2000 and light-chain-specific HRP-conjugated anti-mouse IgG 1:2000 for 1 hour at room temperature.

Gonad dissection and antibody staining

We largely followed protocols for dissecting gonads, fixation, and antibody/DAPI (or Hoechst dye) staining from the laboratory of Dr. Tim Schedl (Washington University in St. Louis School of Medicine), using the methanol/formaldehyde fix for 10-15 minutes. We modified these protocols by manipulating worms in low retention microfuge tubes (instead of glass tubes) and then washing/ blocking the dissected gonads in 4x15minute washes in PBST + 0.1% BSA post fixation. Primary and secondary antibodies were diluted in PBS + 0.1% BSA, and we included two extra wash steps post secondary antibody incubation. Chicken anti-GLD-1 was used at 1:4000,

and Alexa 488 fluorescently-conjugated secondary antibody (Molecular Probes-Invitrogen) was used at 1:2000 dilution. All dissected gonads were blocked, incubated in primary and secondary antibodies, and washed simultaneously in the same conditions.

qRT-PCR of GLD-1-associated messenger RNAs

RNA from anti-GLD-1 and anti-IgY immunoprecipitations and from total input RNA for each biological replicate was reverse transcribed using SuperScript III (Invitrogen) according to manufacturer's instructions. Specifically, we used 0.5µg poly-dT oligonucleotides per reaction, 50°C incubation temperature for 1 hour, and we scaled up the reaction to 50ul. We then used 0.5µl cDNA as template with the Light Cycler 480 SYBR Green I kit (Roche) for quantitative real time PCR (qRT-PCR) according to manufacturer's instructions. We performed negative control reactions with no template for each primer pair to check for nucleic acid contamination.

Data were collected from a Roche Light Cycler 480 machine using manufacturer's software. After scrutinizing melting curve analyses to ensure amplification of a single appropriate product, we imported the data (fluorescence measurements per cycle) into the program LinRegPCR for further analysis (Ramakers et al., 2003; Ruijter et al., 2009). LinRegPCR performs a linear regression-based baseline correction on each sample individually and then identifies the cycles during which each sample was log-linearly amplified (its "window-of-linearity"). It uses linear regression to fit a straight line through each window-of-linearity, and the slope of this line is the log of that sample's amplification efficiency. The software averages the efficiencies of individual samples across primer pairs, creating "amplicon groups", and the average efficiency and C_t value per sample computed by LinRegPCR are used to compute the starting amount of material in each sample in arbitrary fluorescent units.

To make expression comparisons for each gene, we divided the inferred amount of starting material computed by LinRegPCR in the anti-GLD-1 IPs by the starting amount in the mock anti-IgY IPs for each biological replicate. Ratios were averaged across at least three biological replicates and SEMs computed.

RNAi of putative Cbr-GLD-1 targets

We used the WormBase *C. briggsae* gene predictions (build CB3) and the NCBI primer designing tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast) to identity primers that yielded unique *C. briggsae* PCR products 400-900bp in length. Using versions of the primers containing the T7 phage RNA polymerase promoter, we amplified PCR products from either *C. briggsae* genomic DNA or cDNA. Unique, appropriately-sized products were used as template directly for *in vitro* transcription with the MegaScript T7 RNA kit (Ambion) to produce double stranded RNA according to manufacturer's instructions. RNA was phenol:chloroform extracted and ethanol precipitated, resuspended in TE, and run out on an agarose gel to check integrity.

We initially injected dsRNA for two different genes at a time into the gut of adult *C. briggsae* hermaphrodites grown at 20°C. Injected animals were recovered to 20°C, moved to a fresh plate in ~12 hours, and their progeny scored 3-5 days later with both the dissecting and DIC microscopes for mutant phenotypes.

Phylogenetics

We used all 467 amino acids of the Cbr-GLD-1 coding sequence to search the NCBI Protein Reference Sequences database in fall 2009 using BLASTP 2.2.20 with default parameters. We chose hits in choanoflagellates, cnidarians and bilaterians with e-values ~ < $1x10^{-10}$, and we eliminated all partial sequences (identified by lack of initial methionine or short sequences < 100

amino acids), eliminated duplicate or nearly identical sequences (which, for instance, might be the result of sequencing errors or minor alternative splicing), and used only representative metazoan taxa (for instance, we used only *D. melanogaster* sequences from *Drosophila*, and only mice and primate sequences from mammals as many mammalian orthologs had 100% identity to one another.) These filters resulted in 102 sequences.

We used ClustalX 2.0.11 to align these amino acid sequences (Larkin et al. 2007). We performed several alignments, trimming sequences N-terminal and C-terminal to the conserved KH and Qua 1 and 2 domains, though we only trimmed from one terminus at a time in order to more carefully evaluate how trimming affected each iteration of the alignment. We selected the ClustalX option *Iterate each alignment step*, which significantly improved the alignments, and we removed all gaps before each re-alignment. We also changed the Pairwise Alignment gap opening penalty to 35 (from 10) and the gap extension penalty to 0.75 (from 0.1). Each round of alignment resulted in larger blocks of conservation and a shorter overall alignment length; the final alignment length was 451 amino acids.

We used Mr. Bayes (Huelsenbeck and Ronquist, 2003; Ronquist and Huelsenbeck, 2005) to construct a Bayesian phylogeny with these parameters: 1 million generations, burn-in period of 25% of generations, rate parameter set to *adgamma* for the auto-correlated gamma model (in which rates vary across sites according to the gamma distribution but the rate at each site depends in part on the rates at adjacent sites), and the amino acid model *mixed* (which allows the MCMC sampler to explore all the fixed-rate amino acid evolution models and determine how each model contributes to the phylogeny in proportion to its posterior probability). All other run parameters were default settings (including 2 runs of 4 chains each and 1 cold chain in each run). We examined output in order to determine that our cold chains were reasonably mobile, that the standard deviation of split frequencies descended to 0.01, and that the

potential scale reduction factor (PSRF) was near 1 for all parameters and for partition branch lengths. Additionally, we checked the Mr. Bayes-generated plot of *generation vs. log likelihood* to assess whether we had reached stationarity. We gratefully acknowledge the computing resources offered by Adam Bazinet (laboratory of Michael Cummings, University of Maryland) to run Mr. Bayes. Phylogenetic trees were edited using Dendroscope (Daniel et al. 2007).

RESULTS

Creation of an anti- Cbr-GLD-1 *antibody*

In order to identify the messenger RNA targets of *C. briggsae* GLD-1, we developed an anti-Cbr-GLD-1 antibody for use in *C. briggsae* immunoprecipitation experiments. We additionally desired that this antibody recognize GLD-1 in other *Caenorhabditis* species to maximize its utility. To this end, we aligned GLD-1 orthologs in *C. briggsae, C. elegans*, and *C. remanei* and identified stretches of amino acids that were common to GLD-1 but not other homologs. This work resulted in an affinity-purified polyclonal chicken anti-GLD-1 antibody designed to amino acids 16-36 of Cbr-GLD-1. In immunoblots, chicken anti-GLD-1 recognizes a band at the predicted size of GLD-1 in *C. briggsae* as well as in every *Caenorhabditis* species tested: *C. elegans*, *C. remanei*, *C. brenneri*, *C. japonica*, and the unpublished *C.* sp. 5 and *C.* sp. 9. The antibody fails to recognize that band in the *C. elegans* null mutant *gld-1*(*q485*) and *C. briggsae* deletion mutant *gld*-1*(nm68)* (dissertation Chapter 1, Figure 5).

Chicken anti-GLD-1 also recognizes native Cbr-GLD-1. We obtained the same expression pattern in dissected, fixed gonads as reported by Nayak et al. 2005 for *C. briggsae* GLD-1 (Figure 2; Qinwen Liu, unpublished data) and *C. elegans* GLD-1 (Jones et al. 1996): obvious protein expression begins in the middle of the germline mitotic zone, increases in intensity to the

pachytene region, and then abruptly terminates as germ cells enter diplotene of meiosis I.

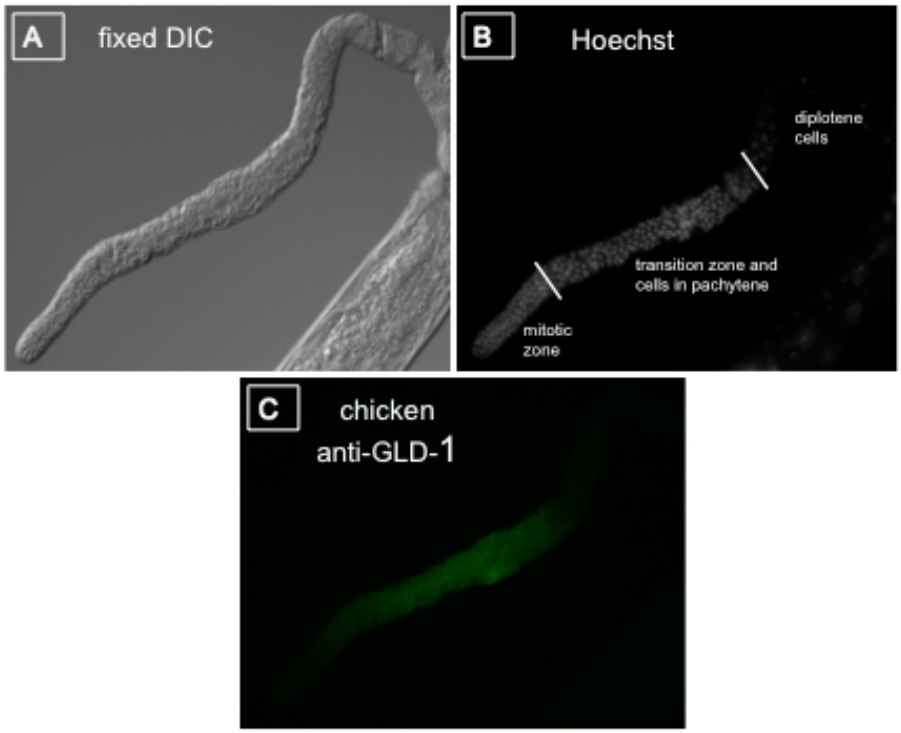


FIGURE 2. Images from Qinwen Liu (unpublished results). Chicken anti-GLD-1 antibody recognizes native GLD-1. Fixed, dissected gonad of a *C. briggsae* wild-type hermaphrodite shown (A) with differential interference microscopy (DIC), (B) Hoechst-stained to observe DNA morphology, and (C) stained with chicken anti-GLD-1. The Cbr-GLD-1 pattern here is the same as observed by Nayak et al. (2005).

Identification of Cbr-GLD-1 mRNA targets with microarrays

Previous studies have identified specific messenger RNA targets of STAR proteins on an individual or small scale (for instance, Lee and Schedl 2001, Wu et al. 2002, Matter et al. 2002, Nabel-Rosen et al. 2002, Di Fruscio et al. 2003, Mootz et al. 2004), or have used *in vitro* selection assays followed by *in silico* searches to predict RNA recognition sites and targets of STAR proteins (Ryder et al. 2004, Galarneau and Richard 2005, Tremblay and Richard 2006, Chawla et al. 2009, Galarneau et al. 2009). We took an *in vivo* genome-wide approach, using RNA immunoprecipitation coupled with microarray analysis ("RIP- Chip"; Tenenbaum et al. 2000) to identify potentially all targets of *C. briggsae* GLD-1 in adult hermaphrodites (see Figure 3).

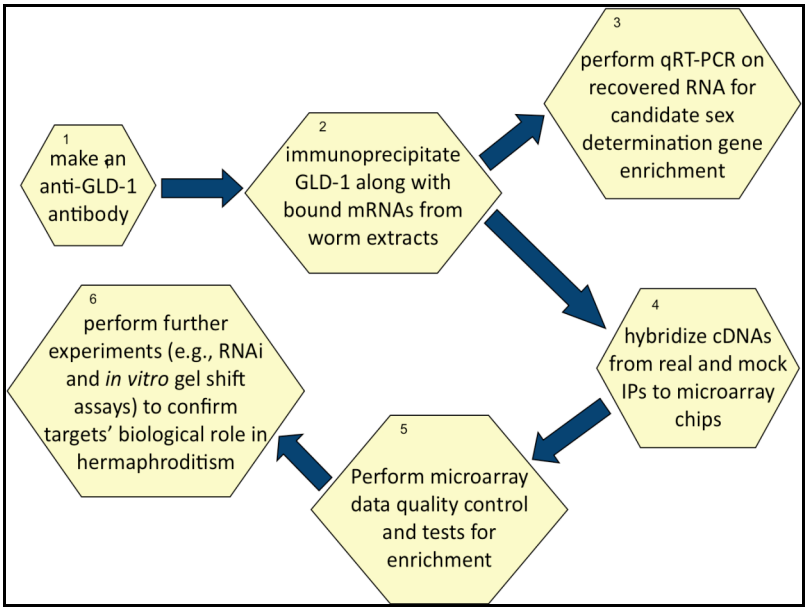


FIGURE 3. Work flow for identifying Cbr-GLD-1 mRNA targets by RIP-Chip.

We immuoprecipitated GLD-1-associated mRNAs from lysates of *C. briggsae* adult hermaphrodites using the chicken anti-GLD-1 antibody described above. Obtaining high quality RNA is key to the subsequent microarray analysis, and Figure 4 reveals the total RNA we recovered by agarose gel electrophoresis from five biological replicataes.

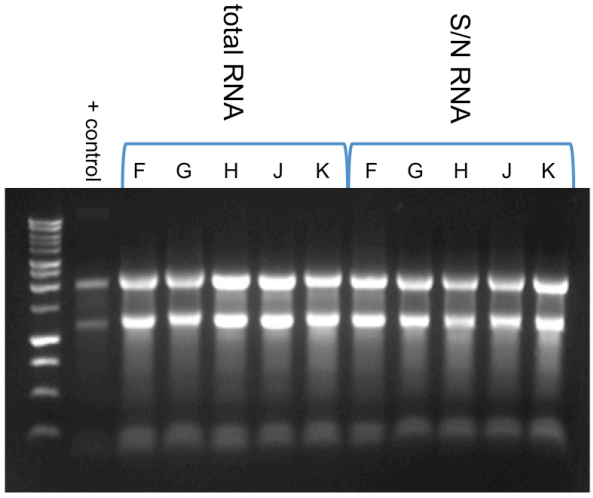


FIGURE 4. 1% TBE-agarose gel electrophoresis of total RNA obtained from worm lysis/pre-clearing steps ("total RNA") and after immunoprecipitation ("S/N RNA") in the RIP-Chip protocol. Each letter (F-K) designates a biological replicate. The bright RNA bands are ribosomal RNA, and their integrity indicates a general lack of RNA degradation.

To check effectiveness of the immunoprecipitations, we performed Western blot analysis on GLD-1 recovered from total input material, post-immunoprecipitation supernatant, and antibody-bound beads using an anti-GLD-1 antibody different than that used for the immunoprecipitations (Jones et al. 1996; Figure 5). We also checked for specificity of the immunoprecipitations with quantitative real time PCR (qRT-PCR) using primers for positive control genes *Cbr-oma* (the single ortholog of the *oma-1* and *oma-2* paralogs in *C. elegans*) and *Cbr-rme-2*, and negative control genes *Cbr-nol-1* and a pan-actin primer set (Lee and Schedl 2004, Nayak et al. 2005; Figure 8, described below).

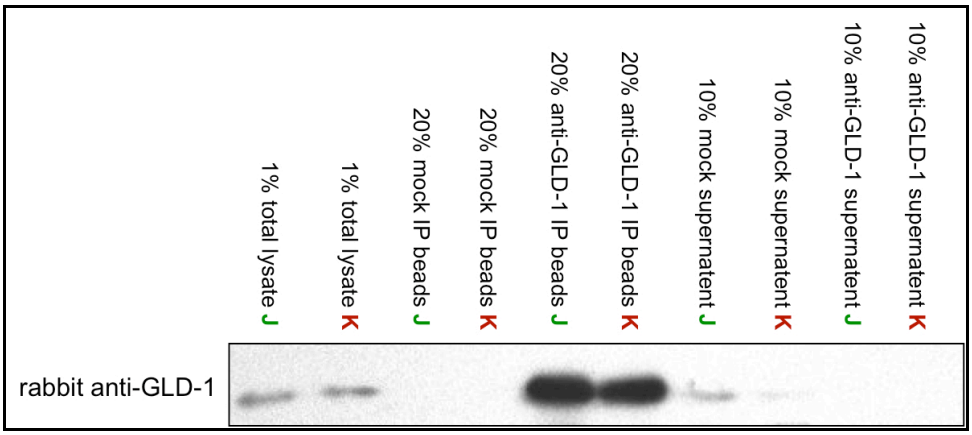


FIGURE 5. Western blot of immunoprecipitation material in two biological replicates. Samples were taken from pre-cleared input material before immunoprecipitation ("total lysate"), from post-incubation antibody-bound beads, and from immunoprecipitation supernatants after incubation. The anti-GLD-1 antibody used in this assay recognizes different epitopes from that used for the immunoprecipitations. "Mock" refers to total IgY-bound agarose beads.

We performed the microarray analysis with a total of 8 two-color arrays: 5 arrays for anti-GLD-1 IP vs. anti-IgY IP mRNA expression comparison and 3 arrays for anti-GLD-1 IP vs. total input mRNA expression comparison. Figure 1 depicts the microarray experimental design, incorporating both biological replication and dye swaps.

RIP-Chip microarray data are non-standard in that even transformed, processed intensity measurements have skewed distributions (for instance, see box plot Figure 6). In our work, this is expected because almost all mRNAs in *C. briggsae* are not bound by GLD-1, nor do they bind to total IgY. Thus the probes for most genes will have low fluorescent measurements, resulting in a great leftward skew in plots of frequency vs. intensity measurement, for instance. Moreover, the signal from true Cbr-GLD-1 targets will increase as the abundance of those target mRNAs in the immunoprecipitations increase, while the signal for the same probes in the anti-IgY IP channels should remain at baseline levels. Accordingly, we also detect an intensity-dependent bias in the data, and anti-GLD-1 IP mRNA intensity measurements are roughly two-fold higher than those for anti-IgY IP immunoprecipitations (Figure 6). This intensity skew and bias is present in the total input mRNA vs. anti-GLD-1 IP mRNA expression comparison as well, though to a lesser extent.

Such characteristics render RIP-Chip data unsuitable for standard microarray normalization methods like loess or quantile normalization, which assume that measurements from the two channels of a dual-color array have nearly the same distributions, that most genes are not differentially expressed between conditions, and that intensity-dependent biases are technical artifacts (Russell S., ed. 2009).

To help ensure appropriate analysis of the data, we performed a 2 x 2 matrix of data manipulations on raw intensity values from both expression comparisons, using 2 different methods to correct for background intensity (no background subtraction and *normexp* as implemented in *limma* (Ritchie et al. 2007, Silver et al. 2009)) and 2 different normalization methods (median scaling and eCADS (Dabney and Storey 2007a and 2007b)). We chose the R-based eCADS software because it uses dye-swaps in a valuable way, combining intensity measurements for probes with a particular treatment over all arrays regardless of what channel

those probes are in, without striving for equal intensities in the two channels or enforcing a particular distribution of measurements.

To assess which set of background-correction/normalization methods performed best for each expression comparison, we created box plots of measurement intensities for each of the data manipulation combinations. We determined that the most uniform box plots for the anti-GLD-1 IP vs. anti-IgY IP mRNA expression comparison resulted from *normexp* background-correction and eCADS normalization (Figure 6 (A)), and for the anti-GLD-1 IP vs. total mRNA comparison, no background correction and eCADS normalization (Figure 6 (B)).

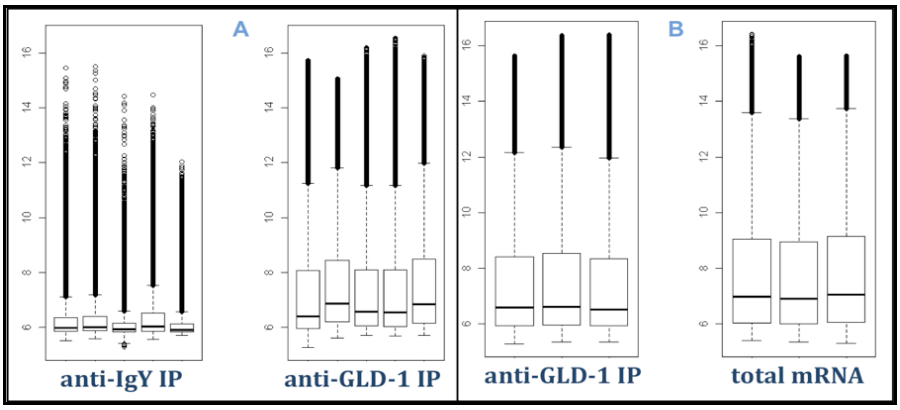


FIGURE 6. Individual array boxplots of (A) *normexp* background corrected, eCADS normalized intensities for the anti-GLD-1 IP vs. anti-IgY IP microarray comparison, and (B) no background-corrected, eCADS normalized intensities for the anti-GLD-1 IP vs. total mRNA comparison.

Next, as the data are not normally distributed (Figure 6), we detected differential gene expression using two non-parametric analysis programs with different mathematical methodologies: SAM (Tusher et al. 2001) and EDGE (Leek et al. 2006, Storey et al. 2007). Both methods compute a differential gene expression statistic (the "d-statistic" for SAM and an optimal discovery procedure (ODP) score for EDGE) and then permute the data in order to assess statistical significance.

Consistent with the hyper-sensitive differences in signal intensity one might expect in a mock vs. real antibody RIP-chip experiment, we found a large list of probes significantly enriched

in the anti-GLD-1 IP vs. anti-IgY IP mRNA expression comparison using SAM (>20,000 with a false discovery rate (FDR) = 0); we found many fewer probes significantly enriched in the anti-GLD-1 IP vs. total mRNA comparison (4,035 with an FDR <2%). Using the more liberal EDGE software, we found even more significant probes for each comparison with FDRs <1%. To conservatively identify genes enriched in the anti-GLD-1 immunoprecipitations, and to guard against method-dependent positive results, we created a Venn diagram of positively enriched probes for both expression comparisons from both differential gene detection programs in order to identify probes in common between all four positive probe lists (Figure 7).

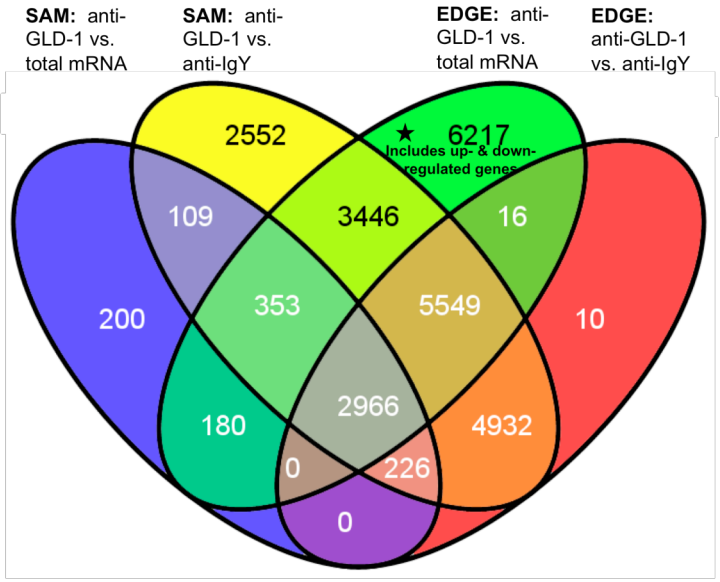


FIGURE 7. Venn diagram (Oliveros 2007) of two microarray comparisons, anti-GLD-1 IP mRNA vs. mock anti-IgY IP mRNA and anti-GLD-1 IP mRNA vs. total input mRNA, each analyzed with two differential gene expression programs, SAM and EDGE. Values in each oval are the number of probes enriched in anti-GLD-1 immunoprecipitations with FDRs of at most <2% (except see * in diagram). 2,966 probes were found in common to all four data sets, representing 802 *C. briggsae* protein-coding genes.

The overlap of these probe lists results in 2966 probes. Lastly, we filtered out genes only represented by a single probe, to eliminate potential false positives, and collapsed the probes

into genes (Cristel Thomas, unpublished Perl script). This results in a list of 802 *C. briggsae* predicted protein coding genes found significantly enriched in anti-Cbr-GLD-1 immunoprecipitations compared to both mock anti-IgY immunoprecipitations and total input mRNA controls with FDRs of at least <2%.

Cbr-GLD-1 does not associate with candidate sex determination genes

To identify potential sex determination targets of Cbr-GLD-1, we interrogated the list of putative GLD-1 mRNA targets from our microarray analysis for candidate sperm-promoting genes. The loss of these candidate genes yields a mutant feminized *C. briggsae* germline phenotype: *she-1* (Guo et al. 2009); *trr-1* (Ron Ellis, personal communication); *fog-3* (Chen et al. 2001); *fog-1* (Ron Ellis, personal communication); *her-1* (Streit et al. 1999); *fem-1, fem-2, and fem-3* (Hill et al. 2006), *puf-2* and *puf-12* (Qinwen Liu, personal communication), and *tra-1* (Kelleher et al. 2008). However, we found that none of these genes appear of the list of 802 target genes.

To confirm these observations, we probed RNAs recovered the from *C. briggsae* anti-GLD-1 and mock anti-IgY immunoprecipitations directly with quantitative RT-PCR for these sperm-promoting genes. We also amplified positive control (*Cbr-oma-2* and *Cbr-rme-2* (Lee and Schedl 2001, Nayak et al. 2005)) and negative control (*nol-1* and pan-actin) transcripts. As shown in Figure 8, we find that positive and negative control genes behave as expected, detecting *Cbr-oma-2* and *Cbr-rme-2* more than 25-fold greater than *Cbr-nol-1* and actin in anti-GLD-1 IP versus mock IP RNA. However, we find that no candidate sperm-promoting Cbr-GLD-1 target sex determination gene is recovered at an appreciable level above negative controls (see below for *Cbr-tra-2*). This includes *Cbr-fog-3*, for which we were able to amplify message from total input

mRNA but not from either anti-GLD-1 IP or anti-IgY IP RNA, likely revealing that *fog-3* transcript levels are too low to be amplified from immunoprecipitation material.

We also interrogated the Cbr-GLD-1 target microarray list for genes known to act in sex determination in either *C. briggsae* or *C. elegans*, regardless of their mutant phenotype. Of the 30 genes we checked for among the 802 putative targets, including *gld-1* itself, only *Cbr-nos-2* appears as a putative Cbr-GLD-1 target. Additionally, Lee and Schedl (2001) showed that in *C. elegans*, the mRNAs for *puf-5, puf-6, and puf-7* (of the PUF family of conserved RNA-binding proteins) are bound by GLD-1. In *C. briggsae*, we do not find any genes in this particular Puf clade among Cbr-GLD-1 targets, but we do find that *Cbr-puf-8* and two PUF genes in the *puf-4* clade (Qinwen Liu, unpublished results) are highly enriched in the target list.

As shown in Figure 8, it is possible that *Cbr-tra-2* is slightly enriched in anti-GLD-1 IPs compared to mock IPs. *tra-2* is a female-promoting gene in *C. elegans* and *C. briggsae* (Kuwabara and Kimble 1995, Kelleher et al. 2008), and *tra-2* is a confirmed target of *C. elegans* GLD-1 (Jan et al. 1997). Even though *Cbr-tra-2* repression by GLD-1 would presumably feminize, not masculinize, the *C. briggsae* germline (yielding a phenotype inconsistent with Cbr-GLD-1 single mutants; dissertation Chapter 1), GLD-1 might still bind to the *tra-2* mRNA just like *C. elegans*, though perhaps with no sex determination consequence. We find by Mann Whitney U test that the difference in fold-enrichment between *Cbr-tra-2* and actin is not significant ($p=0.45$), whereas the difference between *Cbr-tra-2* and *Cbr-oma* is significant ($p=< 0.1$). Especially when considering the overall transcript abundance of each gene in wild-type animals (data not shown), these results indicate that *tra-2* may not be bound by Cbr-GLD-1 in *C. briggsae* as it is in *C. elegans*.

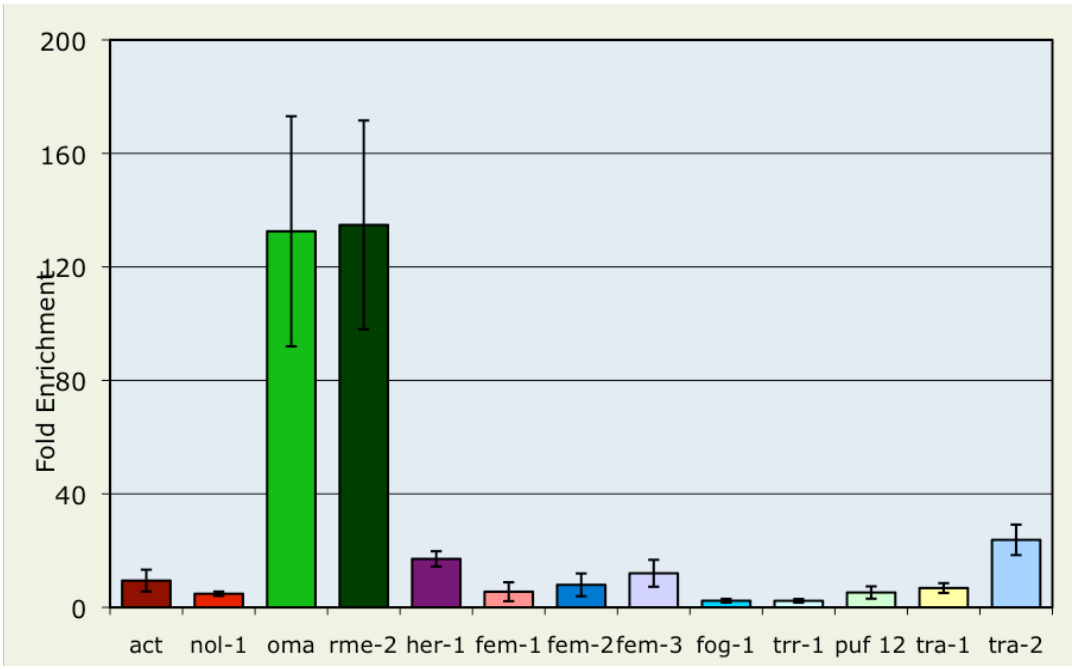


FIGURE 8. Quantitative RT-PCR assessment of Cbr-GLD-1 immunoprecipitated candidate sex determination targets. Enrichment is calculated by dividing the inferred amount of starting material in anti-GLD-1 IPs by that in mock anti-IgY IPs for each gene, averaged over at least 3 biological replicates. Additionally, though we could amplify *Cbr-fog-3* from total mRNA material, we could not detect it in either mRNA recovered from anti-GLD-1 IPs nor from mock anti-IgY IPs for four replicates.

Characterization of Cbr-GLD-1 putative target phenotypes by RNAi

To identify putative GLD-1 mRNA targets that might be involved in *C. briggsae* sex determination, we used RNAi to knock down 100 genes from the microarray target list. These 100 genes had individual probes that received the highest SAM d-statistic score or had the highest fold change in either the anti-GLD-1 IP vs. anti-IgY IP mRNA or the anti-GLD-1 IP vs. total mRNA expression comparisons. These 100 genes also have *C. elegans* homologs with biochemical activity and biological functions that could be compatible with a sex determination role acting through GLD-1 (e.g., they are expressed intracellularly and are not nucleic acid synthesis enzymes).

We injected pairs of double stranded RNA (Gönczy et al. 2000) into the gut of adult *C. briggsae* hermaphrodites, injecting together paralogs and member of gene families together if identified, and separating gene products known to genetically interact with one another. We scored F1 adult progeny 3-5 days later at 20°C with both the stereoscope and DIC compound microscope for mutant phenotypes. When we recovered maternal sterile, embryonic lethal, or sex determination mutant phenotypes, we repeated the injections but used RNA for individual genes.

Appendix 1 lists the genes affected with RNAi in this assay and the germline phenotypes recovered. We did not find any genes that gave clear mutant germline sex determination phenotypes, i.e., all sperm or all oocytes in the hermaphrodite germline. The most common phenotype recovered was defective oogenesis, in which oocytes had abnormal appearances (small and/or unusually shaped and/or with unusual textures). In a few cases, we found preferential 'disintegration' of the most proximal oocyte(s) as well. In most germlines with morphologically aberrant oocytes, we observed sperm "stuck" in the uterus, not in the spermathecae. Other germline mutant phenotypes recovered include reduced germ cell number, slow ovulation/fertilization/embryo laying, the laying of oocytes (not embryos), decaying germ cells (either in the proximal and/or distal germline), and abnormal early cell divisions in F2 embryos. Interestingly, no gene we subjected to RNAi had a phenotype more penetrant than ~33%.

GLD-1 belongs to a nematode-specific clade of STAR-domain proteins

GLD-1 belongs to the STAR family of RNA-binding proteins. STAR proteins are characterized by a single KH domain flanked by two evolutionarily conserved domains

designated Qua 1 and Qua 2 (after the homolog Quaking first described in mice; Ebersole et al. 1996; Figure 9).

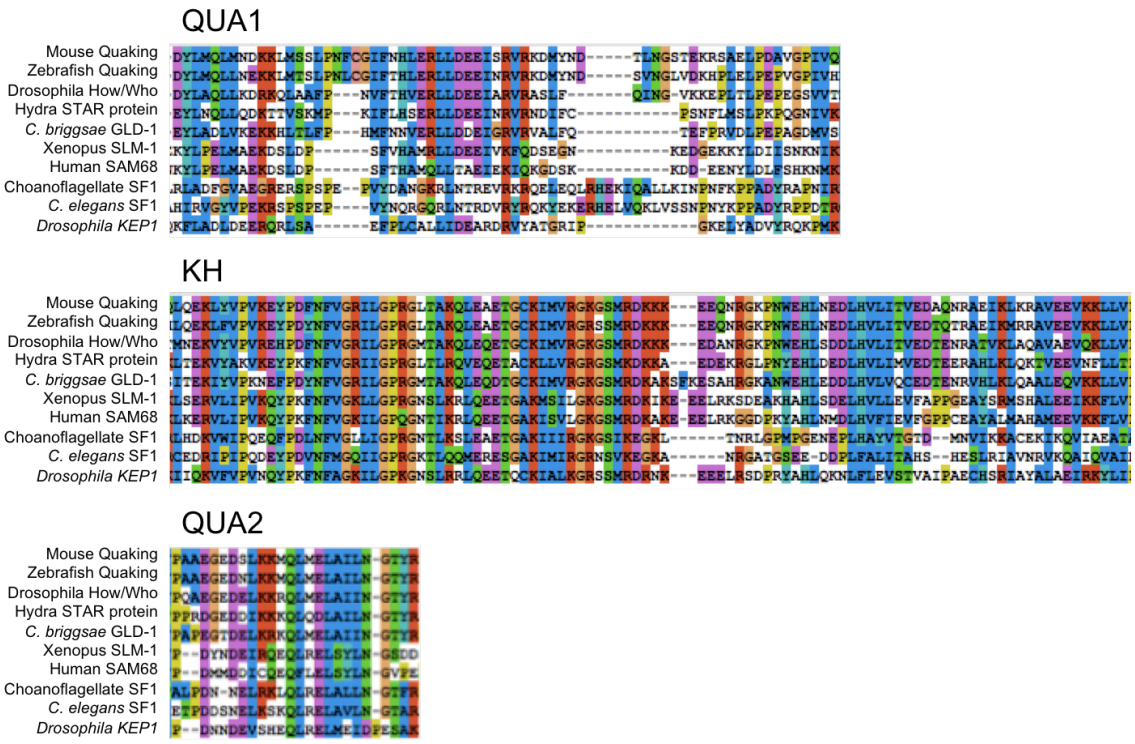
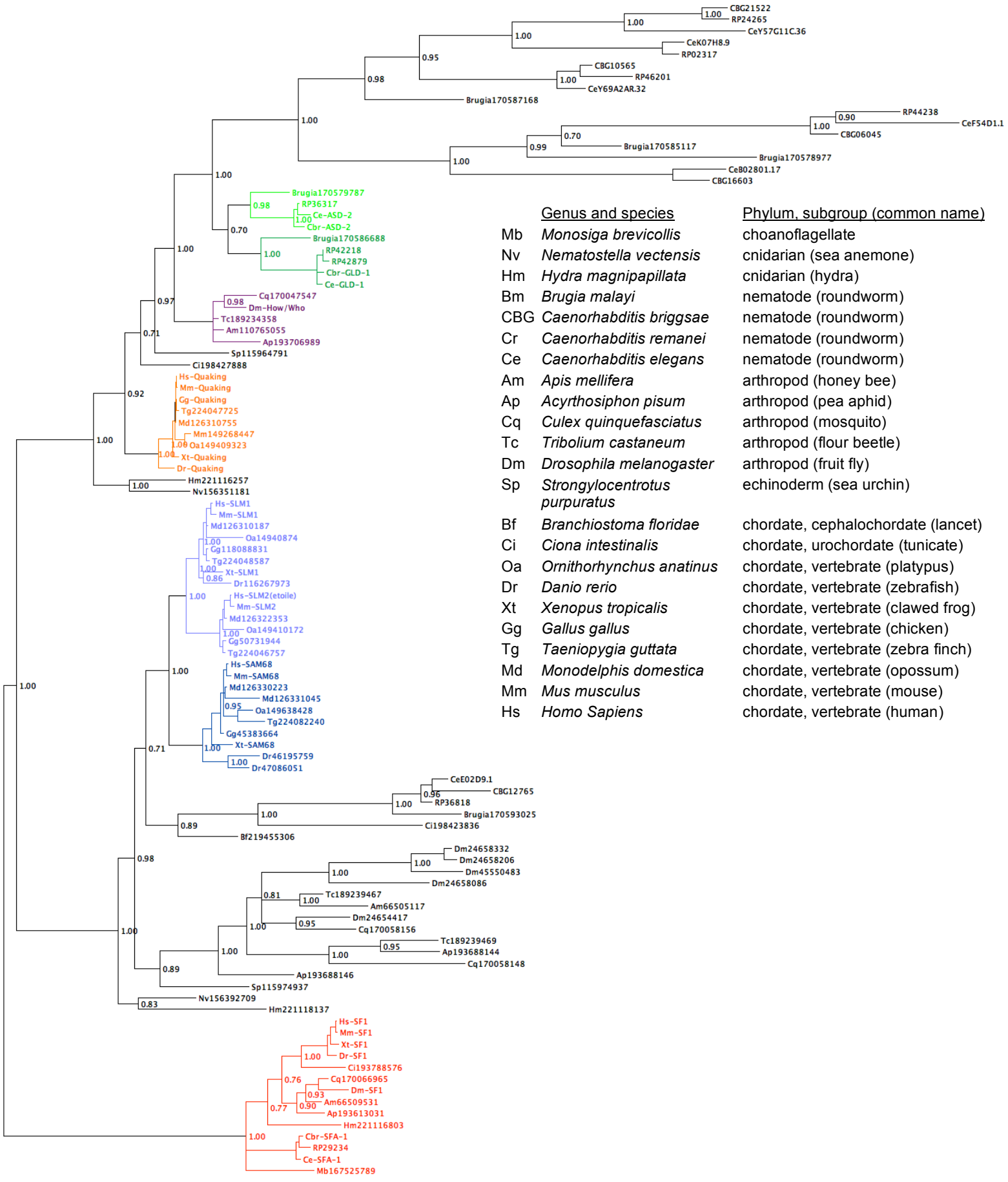


FIGURE 9. Alignment of representative metazoan STAR domains. The three conserved regions are shown in sequence order: QUA1 (involved in homodimerization, except in the SF1 subfamily that remains as monomers; their sequences are more divergent here) (Zorn and Kreig 1997, Chen et al. 1997, Liu et al. 2001, Beuck et al. 2010), the KH RNA-binding domain, and the QUA2 domain (which provides at least an extended RNA-binding surface) (Liu et al. 2001, Ryder et al. 2004, Maguire et al. 2005). In addition, STAR proteins form protein contacts with different binding partners (for instance, Taylor et al. 1994, Clifford et al. 2000, Selenko et al. 2003, Najib 2005, Robard et al. 2006). Alignment was performed with ClustalX 2.0.12 according to default parameters; residue are colored according to ClustalX defaults to highlight chemically similar amino acids.

We desired to place Cbr-GLD-1 in its larger evolutionary context and thus constructed a phylogeny of GLD-1 homologs across metazoans. We took all 467 amino acids of the Cbr-GLD-1 coding sequence and used BLAST against the NCBI Protein Reference Sequences database. We identified 102 seemingly full-length, non-redundant proteins with e-values ~< 10^{-10} from representative metazoa and the choanoflagellate *Monosiga brevicollis*. We trimmed and aligned

these sequences, yielding a multiple sequence alignment of 451 amino acids. We used Mr. Bayes to construct a Bayesian phylogeny with a mixed model of amino acid evolution that samples all fixed rate models in Mr. Bayes and the *adgamma* auto-correlated site rate model (Huelsenbeck and Ronquist 2001, Ronquist and Huelsenbeck 2003). We confirmed stationarity, and the resulting tree, rooted at the SF1 clade of ancient splicing factors (see below), is presented in Figure 10 with posterior probabilities given for internal nodes. Clades containing genes that have been subject to intense study are indicated in color (see figure legend).



	Genus and species	Phylum, subgroup (common name)
Mb	*Monosiga brevicollis*	choanoflagellate
Nv	*Nematostella vectensis*	cnidarian (sea anemone)
Hm	*Hydra magnipapillata*	cnidarian (hydra)
Bm	*Brugia malayi*	nematode (roundworm)
CBG	*Caenorhabditis briggsae*	nematode (roundworm)
Cr	*Caenorhabditis remanei*	nematode (roundworm)
Ce	*Caenorhabditis elegans*	nematode (roundworm)
Am	*Apis mellifera*	arthropod (honey bee)
Ap	*Acyrthosiphon pisum*	arthropod (pea aphid)
Cq	*Culex quinquefasciatus*	arthropod (mosquito)
Tc	*Tribolium castaneum*	arthropod (flour beetle)
Dm	*Drosophila melanogaster*	arthropod (fruit fly)
Sp	*Strongylocentrotus purpuratus*	echinoderm (sea urchin)
Bf	*Branchiostoma floridae*	chordate, cephalochordate (lancet)
Ci	*Ciona intestinalis*	chordate, urochordate (tunicate)
Oa	*Ornithorhynchus anatinus*	chordate, vertebrate (platypus)
Dr	*Danio rerio*	chordate, vertebrate (zebrafish)
Xt	*Xenopus tropicalis*	chordate, vertebrate (clawed frog)
Gg	*Gallus gallus*	chordate, vertebrate (chicken)
Tg	*Taeniopygia guttata*	chordate, vertebrate (zebra finch)
Md	*Monodelphis domestica*	chordate, vertebrate (opossum)
Mm	*Mus musculus*	chordate, vertebrate (mouse)
Hs	*Homo Sapiens*	chordate, vertebrate (human)

FIGURE 10. Bayesian phylogenetic tree of the STAR protein family in representative metazoans. The tree is rooted at the SF1 clade of ancient splicing factors. Values along the tree backbone and at intermediate nodes are posterior probabilities; for visual clarity, terminal node posterior probabilities are not given, but values are similarly high to those shown. The subfamilies of well-studied proteins are highlighted in color: GLD-1 and the related ASD-2 in dark green and light green, respectively; How/Who in purple; Quaking in orange; SF1 in red; and SAM68 and the related SLM-1 and -2 in dark blue and light blue, respectively. Proteins are named with their genus/species abbreviation (see key) and then with either their NCBI Protein sequence database name (like "Quaking"), their NCBI GI identification number, or for *Caenorhabditis* sequences, their WormBase protein ID.

We find that the most closely related clade to GLD-1 is the ASD-2 (for *a*lternative *s*plicing *d*efective) family found in nematodes. GLD-1/ASD-2 belongs to a nematode-specific clade of proteins, How/Who is arthropod-specific, and Quaking is vertebrate-specific. There is also a fast-evolving nematode-specific expansion of proteins sister to GLD-1/ASD-2. Interestingly, single proteins from the two cnidarians, Nematostella and Hydra, fall just basal to the GLD-1/How/Quaking clade. Because all major groups of animals represented in the tree possess a GLD-1/How/Quaking subfamily homolog, and because these proteins are more closely related to each other than to other STAR sub-families, we infer that the cnidarian-bilaterian ancestor possessed a GLD-1/How/Quaking-like gene.

Vertebrate SAM68 proteins are well-studied, and though work has demonstrated a role for them in alternative splicing, these proteins can also be found in the cytoplasm, are phosphorylated by different kinases, bind different proteins, and may act on mRNA translation and have other functions as well (for instance, Matter et al. 2002, Li et al. 2002, Taylor 2004, Najib et al. 2005, Paronetta 2006, Chawla 2009). Sister to the vertebrate SAM68 and SAM-like proteins is a largely unstudied clade of arthropod and nematode sequences, to which two cnidarian proteins are just basal. Interestingly, the tree topology reveals that this subfamily underwent particular expansion in the arthropods. The *D. melanogaster* protein KEP1 is in this clade, known to act in alternative splicing and in oogenesis (Di Fruscio et al. 2003, Robard et al.

2006). That cnidarians are sister to this larger SAM68/SAM-like group containing protostome and deuterostome sequences, each phylum more closely related to itself than to others, makes it likely that as for the GLD-1/How/Quaking clade, a metazoan ancestor possessed a SAM68-like protein(s).

The SF1 subfamily, at which we rooted the tree, also possesses representatives of all major taxa, including the sole choanoflagellate STAR protein. This topology is consistent with SF1's role in a conserved cellular process that predates the evolution of animals, pre-mRNA splicing (Collins and Penny 2005). The presence of choanoflagellate SF1 allows us to infer that metazoans evolved from an ancestor possessing an SF1-like gene.

DISCUSSION

Cbr-GLD-1 associates with hundreds of mRNAs

STAR proteins are highly conserved, important regulators of RNA, and much work has been devoted to uncovering their RNA binding specificities, protein structures, and biological roles in both developmental and diseased states. Previous studies have identified individual or small numbers of STAR protein target RNAs, or have used *in vitro* selection assays followed by computer searches to detect potential RNA targets.

In this work, however, we use an *in vivo* genome-wide approach, RIP-chip, to identify potentially all mRNA targets of *C. briggsae* GLD-1, a germline pleiotropic translational regulator orthologous to *C. elegans* GLD-1. We immunoprecipitated endogenous GLD-1 from *C. briggsae* young adult hermaphrodites and used microarrays containing probes for all predicted *C. briggsae* coding genes to compare mRNAs recovered in anti-GLD-1 IPs to those recovered in both mock IPs and total input RNA. We determined the overlap of positively enriched probes

(with FDRs not greater than 2%) from both expression comparisons and identified 802 genes significantly enriched in the anti-GLD-1 immunoprecipitations. Quantitative RT-PCR of positive and negative control genes and Western blot analysis of immunoprecipitation material confirms the specificity of the immunoprecipitations. From the whole genome analysis, we recovered the only mRNA known to bind *C. briggsae* GLD-1, *rme-2*, and also 10 of 13 orthologs/homologs of gene products known to bind *C. elegans* GLD-1: *oma-1/2, glp-1, mes-3, pal-1*, *cpg-1, air-1.1 and -.2, egg-1,* paralogs of *cpg-2* and *gln-5*, and a homolog of *Ce-H02I12.5*. Additionally, we found that 10 of 10 negative control genes for which we did not expect enrichment in anti-GLD-1 IPs were indeed not enriched. Thus, these 802 genes should be highly specific for mRNAs that associate with GLD-1.

Microarray "gene lists" contain some degree of false positives and/or false negatives, and this list of 802 putative Cbr-GLD-1 mRNA targets is no exception. For instance, the RIP-chip procedure allows recovery of indirectly-associated RNAs. As the nematode germline is rich in mRNAs, mRNA-binding proteins, and other ribonucleoproteins (RNPs) with which Cbr-GLD-1 might associate (Lee and Schedl 2006, Nobel et al. 2008), we expect that some of the 802 targets identified might only be indirectly associated with GLD-1. Additionally, as Cbr-GLD-1 is likely a germline-specific protein as in *C. elegans*, it also might adventitiously bind somatic messenger RNAs that it would never encounter in intact worms. These phenomena likely lend some false positives to the Cbr-GLD-1 target list and contribute to weak but consistent signals above that seen in the mock IP. On the other hand, our false discovery rate cutoff (no greater than 2% for any expression comparison) is conservative. We also note that an antibody versus mock immunoprecipitation comparison is a hyper-permissive background in which to detect differential RNA abundance, but demanding that mRNAs recovered from an antibody IP be enriched more than their levels in total worms is a more stringent test. Importantly, our

utilization of endogenous *C. briggsae* GLD-1 allows us to assay truly biologically relevant levels of messenger RNA for potential targets.

RNAi of putative targets suggest Cbr-GLD-1 regulates multiple aspects of oogenesis

We used RNAi to knock down expression of 100 of the 802 putative Cbr-GLD-1 targets. Though we were particularly interested in sex determination mutant phenotypes (i.e., completely masculinized or feminized XX *C. briggsae* germlines), we did not identify any genes within this 100 that gave such phenotypes. Instead, the most common phenotype we recovered was defective oogenesis, where oocytes had abnormal appearances (unusual shapes and/or cytoplasmic compositions). We sometimes found preferential 'disintegration' of the most proximal oocyte, consistent with its unique meiotic maturation status (McCarter 1999). Also consistent with aberrant oogenesis were observations of germlines with slow ovulation/fertilization. As Cbr-GLD-1 is necessary for oogenic meiosis (Nayak et al. 2005), and *C. elegans* GLD-1 is necessary both for progression through oogenesis and for proper oocyte differentiation (Francis et al. 1995a), these RNAi phenotypes serve to corroborate the specificity of the Cbr-GLD-1 targets identified by microarray.

By choosing to RNAi genes with either a high SAM score (which is computed irrespective of absolute expression level) or a high fold change, we sought to sample both genes that were substantially enriched in the IPs and those that may have important roles but with limited expression levels. It is possible that the target(s) of Cbr-GLD-1 involved in sex determination is redundant, so that unless all other relevant genes are knocked out simultaneously, no phenotype will be observed. The sex determination target(s) may also be pleiotropic, such that the loss of function of this gene results in embryonic lethality or a wholly abnormal germline and thus the sex determination phenotype remains invisible. Additionally, while we hypothesized

that Cbr-GLD-1 is acting to repress a sperm-promoting gene in wild-type hermaphrodites whose over-expression is sufficient to cause masculinization, it does not necessarily follow that loss of the target will result in germline feminization. For these reasons, loss of this gene product might not have a scorable sex determination mutant phenotype, and thus we would not identify it in this assay.

Candidate sperm-promoting genes are not targets of Cbr-GLD-1

Both *C. elegans* and *C. briggsae* GLD-1 have multiple functions in the nematode hermaphrodite germline, including a shared role in progression of meiosis in oocyte-fated cells and in sex determination. However, we have previously confirmed with genetic mutations in *Cbr-gld-1* that *gld-1* has an opposite major sex determination role in *C. elegans* and *C. briggsae*: loss of *Ce-gld-1* results in germline feminization, whereas loss of *Cbr-gld-1* results in germline masculinization (Nayak et al. 2005, dissertation Chapter 1). A cross-species rescue experiment, described in Chapter 1, strongly suggests that a change(s) in protein binding partners and/or messenger RNA targets is responsible for the different *gld-1* sex determination phenotypes in *C. briggsae* and *C elegans* rather than a change in GLD-1 function or regulation itself.

In order to directly identify targets of Cbr-GLD-1 involved in sex determination, we queried the microarray target list for *C. briggsae* genes and homologs of *C. elegans* genes known to be involved in sex determination. We also used quantitative RT-PCR to check for enrichment of 9 *C. briggsae* sperm-promoting candidate target genes recovered from anti-GLD-1 immunoprecipitations. However, we find that none of the 9 candidate genes are strongly enriched in anti-GLD-1 IPs compared to negative controls. We also find that of the 30 genes checked for explicitly on the list of putative targets known to be involved in *Caenorhabditis* germline sex determination, we find only *Cbr-nos-2*. Further, when *Cbr-nos-2* is knocked down

with RNA interference in wild-type hermaphrodites, we detect no germline mutant phenotype. Taken together, these results demonstrate that Cbr-GLD-1 may be regulating novel genes in its control of *C. briggsae* hermaphrodite sex determination.

In *C. elegans*, an important sex determination target of GLD-1 is *tra-2* (Goodwin et al 1993, Jan et al. 1999). Ce-GLD-1 allows hermaphrodite spermatogenesis by binding to specific elements in the 3'UTR of the female-promoting *tra-2* mRNA and subjecting it to translational repression. Regulation of *tra-2* may have been key in the evolution of *C. elegans* hermaphroditism, as a protein binding partner of Ce-GLD-1, FOG-2, is absolutely necessary for hermaphrodite (but not male) spermatogenesis, yet interestingly is a *C. elegans*-specific gene (Clifford et al. 2000, Nayak et al. 2005).

tra-2 is also female-promoting in *C. briggsae* (Kuwabara and Kimble 1995, Kelleher et al. 2008). Even though the major mutant sex determination phenotype of *Cbr-gld-1* is germline masculinization, not feminization like in *C. elegans*, Cbr-GLD-1 might still bind to the *Cbr-tra-2* mRNA just as in *C. elegans,* though perhaps with little consequence. However, *Cbr-tra-2's* absence from the Cbr-GLD-1 microarray target list and our failure to find strong enrichment in directed tests with qRT-PCR suggests that *tra-2* is not a target of Cbr-GLD-1. Direct comparison of the association between GLD-1 and *tra-2* mRNA both in *C. elegans* and *C. briggsae* (via an IP approach similar to that described here) is currently underway in collaboration with undergraduate Haag Lab member Dorothy Johnson.

Previous studies have demonstrated that elements in the 3'UTR of *Cbr-tra-2* mRNA can be subject to translational repression (Jan et al. 1997). Why, then, do we not detect binding of *C. briggsae* GLD-1 to *Cbr-tra-2* mRNA? One reason may be that the study of Jan et al. (1997) used only a somatic assay for 3' UTR-mediated repression via the *Cbr-tra-2* mRNA but stopped short of demonstrating germline repression. If Cbr-GLD-1 is a germline-specific protein, as it is in *C.*

elegans, then Cbr-GLD-1 cannot be responsible for the repression detected in somatic reporter gene assays. More broadly, though germline regulation of *C. elegans tra*-2 by has indeed been demonstrated (Goodwin et al., 1993), somatic regulators of *tra-2* mRNA may explain some or all interactions observed in *in vitro* and/or *in vivo* reporter gene assays in *C. elegans* males, *C. briggsae* hermaphrodites, and *C. remanei* females (Goodwin et al. 1993, Jan et al. 1997, Jan et al. 1999, Haag and Kimble 2000). Similarly, *tra-2* mRNA can also be bound by and/or translationally regulated by elements in *Xenopus* unfertilized oocytes, rat kidney fibroblast cells, HeLa cells, and murine Quaking protein ectopically expressed in *C. elegans* (Jan et al. 1997, Jan et al. 1999, Saccomanno et al. 1999, Thompson 2000). Thus, it seems possible that *tra-2* mRNA is able to illicit a conserved regulation, possibly acting through GLD-1 homologs like Quaking in non-nematode species or by one or more of the ~7 GLD-1 homologs found in nematodes.

Clear evolutionary relationships among STAR protein subfamilies

GLD-1 belongs to the STAR family of RNA-binding proteins (Jones et al. 1995, see Vernet and Artzt 1997 for review). STAR proteins are found in diverse organisms and perform a range of biological functions including translational repression, alternative splicing, and RNA nuclear export for processes like cell division, gametogenesis, early and late embryonic development, and apoptosis (see for instance Kramer and Utans 1991, Zaffran et al. 1997, Pilotte et al. 2001, Wu et al. 2002, Nabel-Rosen et al. 2002, Di Fruscio et al. 2003, Taylor et al. 2004, Lee and Schedl 2004, Ohno et al. 2008, Paronetto et al. 2009).

We desired to place Cbr-GLD-1 in its larger evolutionary context and thus constructed a phylogeny of GLD-1 homologs across metazoans. The length of the three evolutionary conserved domains found in STAR proteins, ~200 amino acids, and their high sequence identity allowed us to recover a tree with high posterior probabilities throughout. This phylogeny is not

inclusive of all animal STAR proteins, but contains all seemingly full-length, non-redundant homologs from representative taxa in order to confidently infer the relationships of subfamilies to one another.

We find strong phylogenetic support for different clades of the most studied STAR proteins: vertebrate Quaking, Drosophila How/Who, nematode GLD-1, the splicing factor SF1, and SAM68. Interestingly, we find that Quaking, How/Who, and GLD-1/ASD-2 are closely-related but phyla-specific protein subfamilies. The SAM68/SLM-1,-2/KEP1 clade also shows phyla-specific evolution. Nematode and arthropod proteins within the SAM68-like clade have longer branches than elsewhere in the tree, signifying greater evolutionary divergence here.

In general, however, we find very little gene family expansion or contraction among the species represented here, perhaps indicating tight control of copy number for STAR proteins. With the exception of the splicing-specific SF1 clade, the most common confirmed mechanisms of action of STAR proteins, regulation of alternative splicing and translational repression, are found in both the SAM68-like and the Quaking/How/GLD-1-ASD-2 super-clades. It will be interesting to investigate how flexible RNA binding proteins can be in their mechanism of RNA control, i.e. how quickly individual proteins can alter their function, or instead how quickly proteins of one mechanism can evolve another. Perhaps as long as STAR proteins retain the ability to bind RNA, their protein-protein binding partners can largely determine their functions (Najib et al. 2005).

Choanoflagellates, protists thought to be the outgroup to metazoans (King et al. 2003), possess a single STAR protein: the splicing factor SF1. This is not unexpected, given that pre-mRNA splicing is an ancestral eukaryotic process. Non-bilaterian metazoans, the cnidarians Hydra and Nematostella, however, possess not only SF1, but also proteins that are outgroups to both the SAM68-like and Quaking/How-Who/GLD-1 super-clades. Given these relationships, we

tentatively posit that the ancestor to metazoans possessed only a single STAR protein, SF1, but as early metazoans diversified, the common ancestor to cnidarians and bilaterians experienced gene duplication events resulting in a very small expansion of STAR gene number and subsequent early differentiation of the SAM68-like and Quaking/How/GLD-1-like genes.

This work has allowed us to investigate the role of *C. briggsae* GLD-1 at different biological scales, from individual gene-by-gene analyses (i.e., determining Cbr-GLD-1's ability to stably associate with candidate sex determination mRNAs), to whole genome analysis (microarray-based identification of all Cbr-GLD-1 messenger RNA targets and their confirmation/analysis), to a phylogenetic investigation of GLD-1's place among STAR proteins separated by hundreds of millions of years of animal evolution. As Cbr-GLD-1 is a regulator of hundreds of mRNAs, their identification with RIP-chip methodology allows detection of functions not easily revealed by genetics. Specifically in the nematode germline, a top-down look at GLD-1 targets can shed light on the connection between germline tissue development and cell differentiation. Furthermore, by taking a comparative approach, either locally against *C. elegans* GLD-1, or more broadly across Quaking/How/GLD-1-ASD-2 subfamily homologs, we may compare how, and speculate as to why, Cbr-GLD-1's biological roles, mechanisms of action, binding partners, and targets have evolved.

CHAPTER 3:

Initial characterization of *nm38*, a germline feminizing allele in C. briggsae

ABSTRACT

Sexual determination and differentiation are critical for many aspects of animal biology. Most animal species produce two sexes, males and females, but other species generate hermaphrodites, single organisms that make both sperm and oocytes in one body. The rarest of all mating systems is androdioecy, in which species have two sexes, males and hermaphrodites. Androdioecy is only known for a few groups of animals, and by studying the developmental genetics of androdioecious species, we may be able to infer the molecular changes involved in rare mating system shifts from gonochorism to androdioecy in animals. Within the nematode genus *Caenorhabditis*, there are two reported androdioecious species, *C. briggsae* and *C. elegans*. Phylogenetic and genetic evidence indicates that *C. elegans* and *C. briggsae* independently evolved androdioecy, each from different male/female ancestral species. Because sex determination in the model *C. elegans* is genetically and biochemically well-understood, we can use *C. elegans* to help investigate sex determination mechanisms in *C. briggsae* in order to probe the convergent evolution of androdioecy in this genus. In this work, we characterize a *C. briggsae* germline sex determination mutant, *nm38* that eliminates spermatogenesis in *C. briggsae* hermaphrodites. Though the allele partially feminizes males, *nm38* strains can be maintained as male/female. We find that *nm38* is not allelic to other known *C. briggsae* feminizing genes, including homologs of *C. elegans* feminizing genes, and that it can suppress spermatogenesis in two other sex determination mutant alleles. We map *nm38* to the right end of *C. briggsae* chromosome II.

INTRODUCTION

Sexual determination and differentiation are critical for many aspects of animal biology, affecting, for example, organisms' developmental programs, morphology and behavior, and population genetics dynamics. Given limited sampling of taxa, Jarne and Auld (2006) estimate that two-thirds of known animals species are gonochoristic, meaning they produce two distinct sexes (males and females), each with distinct gametes (sperm or oocytes). Other animal species generate hermaphrodites, single organisms that can make both sperm and oocytes in one body that may undergo either self-fertilization or cross-fertilization or both.

The rarest of all mating systems is androdioecy, in which species produce both males and hermaphrodites. Androdioecy is only known for a few groups of animals, such crustaceans of the classes Branchiopoda and Cirripedia, the mangrove rivulus (killifish recently placed in the new genus *Kryptolebias)*, and some nematodes of the family Rhabditidae (see Weeks et al. 2006 for review). Given metazoan phylogenies, these androdioecious groups clearly acquired the ability to produce hermaphrodites independently of one another, from different gonochoristic ancestral species. By studying the developmental genetics of these androdioecious species and their particular ecologies, we may be able to infer the proximate and ultimate causes of these rare mating system shifts from gonochorism to androdioecy.

The nematode genus *Caenorhabditis* provides a good system in which to study these mating system shifts. There are 10 published species within *Caenorhabditis*, and most of them are gonochoristic. However, two are androdioecious: the model species *C. elegans* and satellite model *C. briggsae*. The hermaphrodites of these species are essentially modified females that are able to make some sperm from an ovo-testis before switching to oocyte production.

Surprisingly, phylogenetic and genetic evidence indicates that within the genus, *C. elegans* and *C. briggsae* independently evolved androdioecy from different male/female ancestral species (Kiontke et al. 2004, Nayak et al. 2005, Hill et al. 2006, Guo et al. 2009). This situation is an example of convergent evolution, the process by which independent, and potentially unique, changes in a developmental pathway can lead to the same overt phenotypic result. Because sex determination in the model *C. elegans* is among the best understood of developmental genetic pathways, we can use the functional information of *C. elegans* to help investigate sex determination mechanisms in *Caenorhabditis briggsae*. This allows us to dissect the molecular genetic basis of the convergent evolution of androdioecy, and leads us closer to understanding the repeated mating system shifts from gonochorism to androdioecy in this genus.

In this work, we investigate the independent evolution of androdioecy in *C. briggsae* by characterizing a germline sex determination mutant, *nm38*, that eliminates spermatogenesis in *C. briggsae* hermaphrodites. Though it partially feminizes males, *nm38* strains can be maintained as male/female. We find that *nm38* is not allelic to other known *C. briggsae* feminizing genes or orthologs of *C. elegans* feminizing genes and that it can suppress spermatogenesis in two other sex determination mutant alleles. We map *nm38* to the right end of *C. briggsae* chromosome II.

METHODS

Forward mutagenesis

Forward mutant screens were performed using 50mM EMS on synchronous *C. briggsae* (AF16 strain) young adult hermaphrodites (just after their L4-adult molt) nutating for 4 hours at

room temperature in M9. Mutagenized animals were extensively washed in M9 and plated on standard NMG plates at 20°C. We singled F1 L4 hermaphrodites two to a 6cm plate and let them lay F2 progeny at either 20°C or 25°C (for warm temperature-sensitive allele screens). Germline feminized (Fog) mutants were identified by screening adult F2 with a dissecting microscope for the characteristic "stacking oocyte" Fog phenotype (as in Schedl and Kimble (1998) and Barton and Kimble (1990)). Fogs were picked from individual plates containing at least ~3 Fog-looking animals and were mated individually to wild-type AF16 males to check for propagation of the Fog phenotype. Wild-type-looking siblings of these Fogs were also singled simultaneously to pass the mutant phenotype in case Fog oocytes were not capable of being fertilized. Screening 7,000 haploid genomes this way, we recovered one true Fog allele, *nm38*. This allele was outcrossed at least 6 times prior to characterization to wild-type males. (We also recovered a number of Spe-like alleles. These mutants display a Fog-like phenotype under the dissecting microscope, but when examined at the L4/young adult stage with DIC, they show evidence of sperm production.)

Western blotting

Western blotting was performed according to standard procedures. In particular, fifty animals of a particular genotype or sex were picked into a small amount of PBS in a microfuge tube, washed once in PBS, resuspended with an equal volume of standard 2X protein sample buffer with 5% BME and heated to 95°C for 5 minutes, then frozen at -80°C until use. Anti-RME-2 antibody was a gift of Dr. Barth Grant (Rutgers University) and was used at 1:2000. Anti-MSP was a gift of Dr. David Greenstein (University of Minnesota) and was used at 1:5000. Anti-tubulin antibody (Sigma T9026) was used at 1:2000.

DNA sequencing

We used Big Dye v3.1 Cycle Sequencing chemistry (Applied Biosystems) for DNA sequencing according to the manufacturer's instructions. Sequencing was performed on ABI 3100 or 3750 machines according to standard protocols, and trace files were examined with the "4 Peaks" program (Mekentosj software http://mekentosj.com) and analyzed or further manipulated with Vector NTI software (Invitrogen).

Mapping of nm38

To test for the linkage of *nm38* to *Cbr-fog-1* and *Cbr-fog-3*, we used standard procedures to sequence *Cbr-fog-1* and *Cbr-fog-3* in the *C. briggsae* mapping strain HK104 to find polymorphisms in those genes between HK104 and our mutagenesis strain AF16. To perform the mapping experiments, we mated *nm38* Fogs to HK104 males and then singled their L4 hermaphrodite cross-progeny to new plates. We picked individual F2 Fogs from these plates to single-worm lyse according to standard procedures and then genotyped them at either the Cbr-*fog-1* or Cbr-*fog-3* locus, while also genotyping pure AF16 and HK104 animals and wild-type looking siblings of F2 Fogs as controls. As the polymorphism we assayed in *Cbr-fog-3* was a single-nucleotide polymorphism that produced a restriction-fragment length polymorphism when cut with the restriction enzyme Xba1 (i.e., a "SNP-snip"), we first used PCR to amplify the relevant part of *Cbr-fog-3* in our F2 Fogs and control animals and then digested purified PCR products with XbaI according to manufacturer's protocol (New England Biolabs). For Cbr-*fog-1*, as the polymorphism we assayed was a 12 base pair insertion/deletion difference between AF16 and HK104, we used PCR to amplify the relevant *fog-1* fragment from F2 Fogs and control animals and then directly ran this product out on a 4% percent NuSieve 3:1 agarose gel (Lonza Rockland) to resolve base pair differences.

To test for linkage of *nm38* to particular *C. briggsae* autosomes, we used the *C. briggsae* genetic map (B. Gupta et al. 2007) and mutations *lev-unc(*sy5440) I; *dpy*(nm4), *cby*(sy5148) II; *cby*(s1272) III; and *cby*(sy5027) IV; (there were no phenotypic markers for Chr. V at the time of this work). In particular, we mated XX homozygote phenotype mutants to *nm38*/+ males, singled and selfed F1 hermaphrodite L4 progeny, and then counted the proportion of wild-type, single mutants, and double mutants among the F2 where the mother was a double heterozygote. At least 1100 animals were counted for each mapping cross.

To map *nm38* by bulk segregant analysis using insertion/deletion polymorphisms ("indels") and single nucleotide polymorphisms that create restriction fragment length polymorphisms ("SNP-snips"), we took advantage of polymorphisms detected and made available by Dr. Ray Miller and Dan Kobalt (Washington University in St. Louis, School of Medicine, http://snp.wustl.edu/snp-research/c-briggsae). We obtained F2 Fogs and control animals as described above using HK104 or VT847 males in the parental cross. We made pools of 100 Fogs, 100 wild-type-looking siblings, 100 AF16 hermaphrodites, and 100 VT847 or HK104 hermaphrodites. We lysed these bulk worms to obtain DNA, and then amplified particular polymorphic loci on specific fingerprint contigs using computationally identified primers (Kobaldt 2010). Although we tried to quantify agarose band gel intensity of the mapping and mutagenesis strain PCR products using spot densitometry, in the end we simply judged intensity differences, and thus linkage between *nm38* and specific mapping assays, by eye.

To perform SNP mapping of *nm38* by sequencing, we single-worm lysed individual F2 Fogs and control animals as described above in 20ul lysis buffer. 5ul of each lysed worm was used in a PCR reaction to amplify specific SNPs using primer determined as above. PCR products were check by agarose gel and then were sequenced as described above.

RESULTS

Forward mutagenic screens for Fog alleles

To obtain mutant alleles that feminize the germline in *C. briggsae* (so called Fog mutants, for ***f****eminization* ***o****f* ***g****ermline*), we performed forward mutagenesis screens in AF16 wild-type hermaphrodites according to the scheme in Figure 1. Screening 7,000 haploid genomes, we recovered one true feminizing allele*, nm38*.

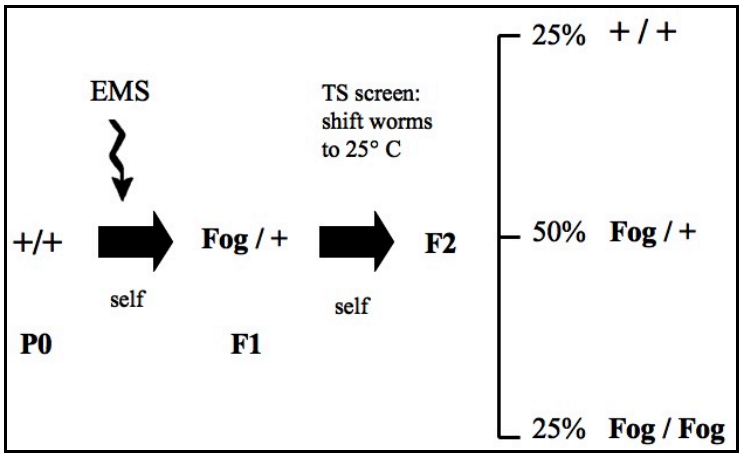


FIGURE 1. Forward mutagenesis scheme. Young adult wild-type *C. briggsae* AF16 hermaphrodites are exposed to 50mM EMS for 4 hours at room temperature. F1 self-progeny are picked to fresh plates and allowed to self. Recessive Fog alleles are recovered in the F2.

nm38 is a germline feminizing allele

nm38 transforms the fate of sperm cells to oocytes in XX and XO animals (Figure 2). Heterozygous *nm38* mothers continuously reared at 20°C produce 20.5 ± 0.9% (n= 8900) Fogs. Normally at the L4 stage, *C. briggsae* XX animals undergo spermatogenesis. However, Figure 2(A) shows the germline of a *nm38* L4 XX worm in which the very first germ cells to differentiate do so as oocytes. When this same animal was left to develop overnight and then examined the next day (B), she had laid no embryo, as a normal adult would have, and has an absence of sperm and presence of the classic Fog "stacking oocyte" phenotype (in which oocyte production is not balanced by embryo laying). This assay distinguishes true Fog alleles from sperm and

fertilization-defective (i.e., Spe and Fer) mutants. We performed this same assay with 23 other XX progeny from a selfing *nm38*/+ mother. In these assays, we found in 7/23 animals with no signs of sperm and 15 animals that possessed sperm/spermatocytes (Figure 2(C), even in the absence of obvious oogenesis. When recovered and examined the next day, the 7 animals that contained no sperm had clearly stacking oocytes and had produced no self-progeny, while the remaining animals were self-fertile.

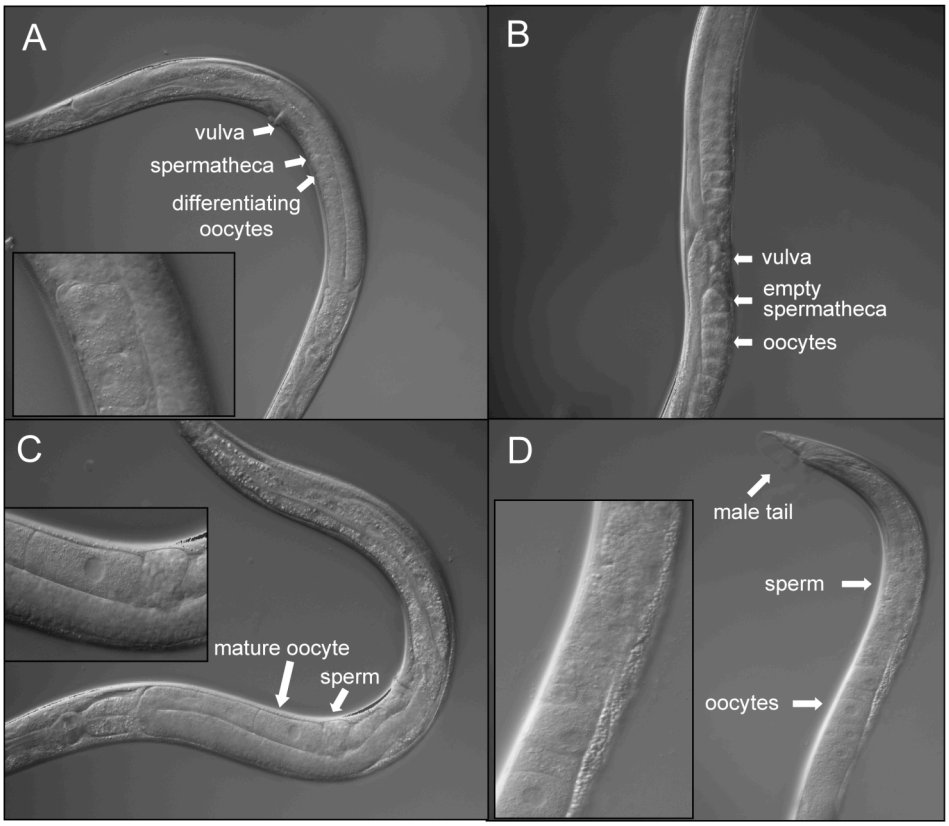


FIGURE 2. *nm38* mutants are Fog. (A) presumed *nm38* XX young adult animal in early gametogenesis. The first germ cells to develop are oocytes, and the inset highlights an empty spermathecae and an oocyte as the most proximal gamete; (B) the same animal as in (A) recovered overnight to a fresh plate. She laid no self-progeny, still possesses no sperm, and displays the stacking oocyte Fog phenotype; (C) young adult wild type-looking sibling of (A) clearly undergoing spermatogenesis; (D) *nm38* XO male with a perfect male soma but both sperm and oocytes in the germline.

nm38 homozygous males first make sperm in their germlines as do wild-type males, but they then switch to producing oocytes a day into adulthood (Figure 2 (D)). Their somatic

anatomy and mating behavior, however, remain normal, and *nm38* XO animals can sire cross-progeny before switching to oogenesis.

We further confirmed *nm38* feminization of the germline with Western blot analysis using a sensitive antibody to the normally abundant Major Sperm Protein (MSP). As shown in Figure 3, Fog XX animals have no detectable MSP compared to the same number of WT young adults.

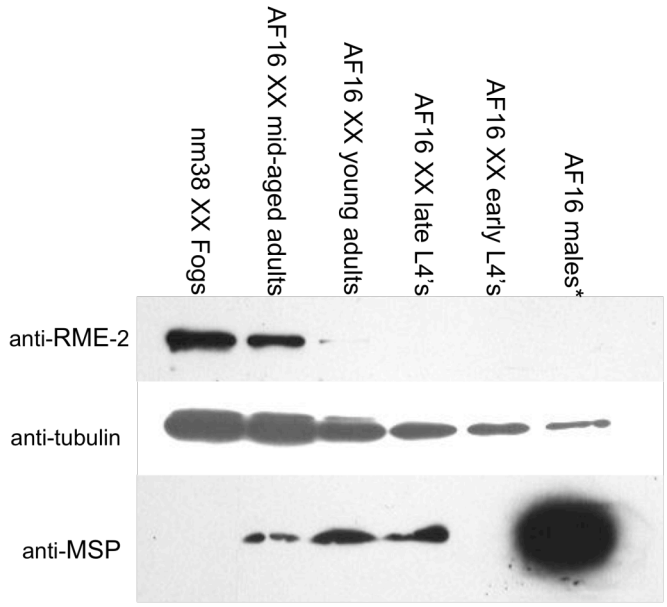


FIGURE 3. Western blot of *C. briggsae* AF16 wild-type and *nm38* mutant animals. Protein in the top panel was exposed to an antibody to the oocyte yolk-receptor RME-2. Protein in the bottom panel was exposed to an antibody to the Major Sperm Protein. Middle panel is tubulin loading control. * Only half as many AF16 males are loaded compared to the XX lanes.

nm38 may be incompletely expressive

As stated above, *nm38*/+ mothers lay only 20.5% Fogs, not the 25% one would expect from Mendelian segregation. This is true despite extensive outcrossing of the allele. When mated into by AF16 males, *nm38* homozygous mothers also lay some inviable embryos. Embryonic lethality is obtained from other combinations of *nm38* crosses as well (e.g., *nm38*/+

hermaphrodite x *nm38* male), and we have not quantified this effect. Thus, the "missing 5% Fogs" may be due to this lethality.

Another explanation for the recovery of only 20%, not 25%, Fogs from a selfing heterozygous mother is that *nm38* is only partially penetrant or expressive, such that the "missing 5%" of Fogs look like wild-type. Indeed, we find that some likely *nm38* XX animals can produce a limited number of Fog self-progeny. We first identified these "self-fertile Fog" animals by singling L4s from an *nm38*/+ selfing mother. To our surprise, we occasionally found Fog-looking worms that produced very few progeny (between 1 and 40 animals), including males, and that these progeny became Fog themselves (including the XO animals, all of which develop oocytes).

To help establish that these "self-fertile Fogs" are *nm38* homozygotes, we mated their Fog self-progeny to wild-type males. We reasoned that if some *nm38* Fogs are capable of limited self-fertility, all of their progeny must also be *nm38*; then when mated into by wild-type males, all F1 cross-progeny will be *nm38*/+ and will be able to produce Fogs in the F2 generation by selfing. We tested this idea by mating into 8 of 20 females that were the progeny of a presumed *nm38* "self-fertile Fog." We singled 5-10 L4 hermaphrodite F1s from each of the 8 matings, let the F1s self, and then phenotyped their F2 self progeny. As shown in Table 1, each mated P_0 Fog produced F1 progeny that all yielded Fogs in the F2 generation. These results strongly suggest that the P_0 Fogs were of the genotype nm38/nm38 and thus are likely the progeny of a weakly self-fertile *nm38* homozygous Fog animal.

Table 1

presumed (m-/-) *nm38* X AF16 matings	F1 plates with Fogs / total plates	% Fogs produced collectively
A	9/9	15.7
B	8/8	15.5
C	8/8	14.6
D	9/9	20.3
E	10/10	16.1
F	5/5	15.0
G	9/9	11.0
H	5/5	14.3

F1 mothers in this assay also produced dead embryos, although we did not quantify this effect. Additionally, the proportion of Fogs recovered in this assay is only 15.5 ± 0.9%, not the expected 20.5% Fogs from typical selfing *nm38*/+ mothers. This difference is significant by unpaired t-test, p-value < 0.005. Most simply, the "lost Fogs" might be accounted for by the dead embryos we observed from F1 mothers. Alternatively, the F1 *nm38*/+ mothers in this assay and typical F1 *nm38*/+ mothers differ in a subtle aspect of their genetic heritage, that is, the maternal *nm38* genotype of the P_0 animal; perhaps a complicated genetic interaction may explain this different Fog phenotype proportion.

Creation of an nm38 homozygous stock

Because *nm38* males are weakly cross-fertile, we were able to establish an *nm38* homozygous stock by letting rare *nm38* males produced from weakly self-fertile *nm38* mothers mate with their *nm38* female siblings. Although, this stock becomes predominantly composed of selfing hermaphrodites at 15°C over time, presumably due to the temperature-sensitivity of *nm38* and rescue of self-fertility at 15°C, at 20°C and 25°C (see below), the *nm38* homozygous stock propagates itself as a male/female strain.

nm38 acts maternally at the L4 stage to produce sperm

We performed a standard temperature-shift experiment to determine the time of *nm38* action. We first grew five *nm38*/+ mothers at 25°C until adulthood. We let them lay for 1.5 days at 25°C and then moved each mother to a fresh plate to lay at 15°C. We counted the number of F1 Fog progeny at adulthood on each plate and found no difference in the proportion of Fogs laid at 25°C vs. 15°C, 23 ± 0.7% and 22 ± 1.7%, respectively (paired t-test p= 0.32). Thus, *nm38* is not a zygotically temperature sensitive mutant allele.

However, when 14 *nm38*/+ mothers were grown at 25°C and either remained at 25°C or were shifted down to 15°C at the last larval stage, L4 (i.e., before the adult molt), we found a significant difference in the percentage of Fog progeny from mothers moved to 15°C at L4 (13.3 ± 1.2% Fogs) compared to those kept at 25°C (21.1 ± 1.1% Fogs) (unpaired t-test p= 0.0004). Taken together, these findings indicate that *nm38* is maternally, but not zygotically, temperature sensitive, and that it seems to require the permissive temperature during the L4 stage to at least rescue some self-fertility of *nm38* homozygotes. Additionally, given our observations of unhatched embryos and rare, developmentally compromised F1 and F2 progeny from weakly self-fertile *nm38* females, we infer that *nm38* also has a role in embryo/larval viability, and that this role is possibly maternal (data not shown).

nm38 produces intersexual somas in certain crosses

We normally propagate *nm38* by mating into mutant females with wild-type males and then letting F1 hermaphrodite heterozygotes produce self-progeny to obtain Fogs in the F2 generation. However, we noticed that during certain kinds of crosses involving *nm38*, we obtained not only XX and XO animals with feminized germlines, but also some animals with intersexual somas (Figure 4). As indicated in Table 2, we find a small percentage of intersexual

somas when an *nm38* heterozygous hermaphrodite mates with an *nm38* homozygous XO animal, or vice versa, and we also see intersexual somas within the *nm38* homozygous stock. However, we detect 10-fold more intersexual somas in heterozygote sib-matings.

Table 2

Maternal genotype	Paternal genotype	Intersex. somas	XO Fogs	XX Fogs
selfing heterozygote		0/>1000	N/A	1825/8900 (20.5%)
Heterozygote	wild-type	0/500	0/50	0/50
Homozygote	wild-type	0/750	0/100	0/50
wild-type	heterozygote	0/300	0/50	0/50
wild-type	homozygote	3/300 (1%)	3/60 (5%)	0/50
Homozygote	heterozygote	6/500 (1%)	89/180 (49%)	83/165 (50%)
Heterozygote	homozygote	9/500 (2%)	44/85 (52%)	48/100 (48%)
Homozygote	homozygote	10/500 (2%)	46/50 (92%)	50/50 (100%)
Heterozygote	heterozygote	114/565 (20%)	57/190 (30%)	62/255 (24%)

All maternal genotypes of XX and XO parents are *nm38/nm38*, except for wild-type animals, who are m(+/+).

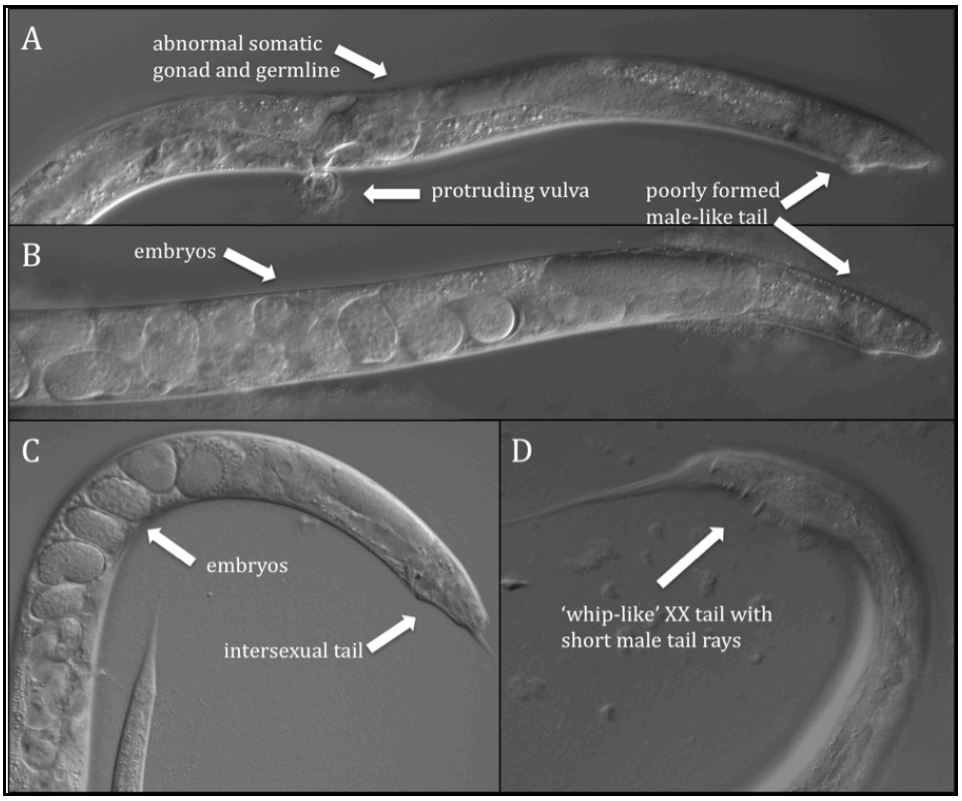


FIGURE 4. *nm38* intersexual somas. (A), (B), (C) are F1 from heterozygote sib-crosses; (D) F1 from AF16 XX x *nm38* XO cross.

Interestingly, we never observe intersexual somas among the progeny of selfing *nm38*/+ hermaphrodites. Thus, these intersexual animals may have an XO karyotype, as they seem to only derive from XX/XO matings. Interestingly, intersexual animals can produce viable self-progeny: Of 20 intersexual animals from which progeny were able to hatch and escape the compromised soma of their parent, we found 32 wild type-looking hermaphrodites, 14 XX Fogs, 2 wild type-looking males, 2 intersexual animals, and many dead embryos. This argues that the intersexual animals are not *nm38* homozygotes, as they are able to produce wild type-looking progeny. It is also unclear why a mating between *nm38* homozygotes produces the severest effect in Table 2. Importantly, there are a significant number of unhatched embryos produced in many of the crosses in Table 2, including the heterozygote x heterozygote cross. We have not

quantified this effect, but its quantitation may shed light upon the true proportion of affected animals and severity of phenotypes in each cross.

Double mutant analysis with nm38

In our efforts to understand the genetic pathway(s) that allows hermaphroditism in *C. briggsae*, it is vital to know where in that pathway each gene product acts. To this end, we performed double mutant analysis with *nm38* and other null or strong hypomorphic *C. briggsae* sex determination mutant alleles with opposite phenotypes.

C. elegans TRA-2 is a transmembrane protein that acts soon after the divergence of the dosage compensation and sex determination genetic pathways to help specify the female soma and germline (Hodgkin and Brenner 1977, Okkema and Kimble 1991, Kuwabara and Kimble 1995). Loss of function mutations in *Cbr-tra-2* masculinize the hermaphrodite soma and germline of *C. briggsae* as they do in *C. elegans* (Kelleher et al. 2008). We used both the *Cbr-tra-2* temperature-sensitive allele *ed23*ts (which largely looks wild-type at 15°C but becomes masculinized at 25°C) and the stronger loss-of-function allele *nm1* in this experiment. Andrew Cheng, an undergraduate student in our laboratory, and I first mated *nm38*/+ males to Cbr-*tra-2*(*ed23*)ts hermaphrodites grown at the permissive temperature or to *nm1*/+ hermaphrodites (at 20°C). We then singled 50 total L4 hermaphrodite progeny from successful matings, let them lay at 25°C, and examined the F2 generation for mothers that had produced both *tra-2*(lf) pseudo-male somas and Fog *nm38* animals. (Because *tra-2* pseudo-males cannot inseminate XX animals, we did not have to separate Fogs from pseudomales in this experiment.) Both Cbr-*tra-2* and *nm38* are on *C. briggsae* linkage group II, and we estimate that the *dpy(nm4)* phenotypic marker, which is very tightly linked to *Cbr-tra-2*, and *nm38* are about 19 map units apart (data presented below). Given this physical linkage, we expected to find *ed23*ts *nm38* or *nm1 nm38*

double recombinants at a frequency of 0.006 among F2 progeny of double heterozygous mothers. Thus anticipating that 9% of animals with Tra somas to also be *nm38* homozygotes, we then examined the germline phenotype of 150 Tra *ed23*(*ts*) and 100 *nm1* animals with Nomarski optics. We found 22/250 Tra animals with clear oocytes (8.8%) (Figure 5). As a control, we scored the germlines of 250 animals with Tra somas from *ed23*/+ or *nm1*/+ mothers who did not produce *nm38* Fogs. Of these 250, none had oocytes in their germlines.

FIGURE 5. Presumed *nm38 ed23*ts double mutant reared at the restrictive temperature with clear oocytes in the germline instead of only sperm.

Thus, if the *ed23*ts or *nm1* and *nm38* alleles act in the same genetic pathway of sex determination and *nm38* is a loss-of-function mutation, we conclude that *nm38* is epistatic to *ed23*ts, such that loss of the *tra-2* gene product does not suppress the production of oocytes in double mutants.

The products of the *C. elegans fem* genes act downstream of *Ce-tra-2*, inhibiting the female-promoting action of TRA-2 to allow male somatic development and sperm production in both males and hermaphrodites (Kimble et al 1984, Hodgkin 1986, Rosenquist and Kimble 1988, Pilgrim et al. 1995, Chin-Sang and Spence 1996, Mehra et al. 1999). In *C. briggsae* however, though the somatic role of the *fem* genes is conserved, they are no longer necessary for the initiation of spermatogenesis (Hill et al. 2006). To determine the epistatic relationship between

nm38 and the *Cbr-fem-3* deletion mutant *nm63*, Shanni Silberberg, an undergraduate student in our laboratory, and I mated *nm38*/+ males to *nm63* homozygous mothers in bulk. These mothers were singled to fresh plates to lay progeny, and from the *fem-3* mothers that produced >40% males (indicative of a successful cross), we singled and selfed 20 of their L4 hermaphrodite progeny. Half of these singled F1s should be of the genotype *nm38*/+; *nm63*/+ and half *+/+; nm63*/+. Of the 20 F1 that we let lay self-progeny, 11 produced Fogs. Finally, we genotyped these F2 Fogs for the *nm63* deletion to ask: can an *nm38 nm63* double homozygote look Fog, due to *nm38* suppression of sperm production in *nm63* hermaphrodites? We found that 12/66 genotyped Fogs were *nm63* homozygotes, 17 were wild-type for the *fem-3* locus, and 37 were heterozygous. By chi-squared test, these observations do not differ significantly from those expected if *nm38* is fully epistatic to *nm63* (X=1.73, df=2). Thus, if *nm38* and Cbr-*fem-3*(*nm63*) are acting in the same genetic pathway and *nm38* is a loss-of-function mutation, we conclude that *nm38* is epistatic to *fem*-3*(nm63)*.

nm38 is not allelic to known Fog genes

Loss-of-function mutations in the *fog-1* and *fog-3* genes in *C. elegans* feminize the germline of both hermaphrodites and males (Barton and Kimble 1990, Ellis and Kimble 1995, Chen et al. 2000, Jin et al. 2001), and RNAi-mediated inhibition or mutation of their *C. briggsae* orthologs does the same (Chen et al. 2001; Dr. Ronald Ellis, personal communication). Thus, *nm38* could be allelic to *Ce-fog-1* or *Ce-fog-3*. To investigate this possibility, we mated *nm38* homozygous females to males of the *C. briggsae* mapping strain HK104. We singled L4 hermaphrodite cross-progeny, let them produce F2 self-progeny, and then genotyped F2 Fogs for polymorphisms we identified in *Cbr-fog-1* and *Cbr-fog-3* between our mutagenesis and mapping strains (see *Methods*). Scoring 40 F2 Fogs for each polymorphism, we found numerous

HK104 heterozygous and homozygous animals, indicating that there is no linkage of *nm38* to either Cbr-*fog-1* or Cbr-*fog-3*.

Other genes in *C. briggsae* are also known to produce Fogs when affected with loss-of-function mutations or RNA interference. *she-1* and *trr-1* were isolated in the laboratory of Dr. Ron Ellis in forward genetic screens for *C. briggsae* feminizing mutations; loss-of-function mutations in *Cbr-she-1* feminize the germlines of hermaphrodites (though they also feminize the germlines of some males late in life), and mutations in *Cbr-trr-1* feminize the germlines of both sexes (Dr. Ron Ellis, personal communication, and Guo et al. 2009). Additionally, our laboratory has demonstrated that knockdown of *Cbr-puf-2* and *Cbr-puf-12* together by RNAi produces Fog animals (Qinwen Liu, unpublished data). *Cbr-trr-1*, *puf-2*, and *puf-12* are all on *C. briggsae* chromosome 2. As we determined that *nm38* is also on chromosome 2 (see below for mapping data), we performed a complementation test of *nm38* with the strong *trr-1* allele *v76* and with a *puf-2* deletion allele, *nm66* (R. Ellis, unpublished data; Q. Liu, unpublished data).

To test complementation with *Cbr-trr-1*, we mated both *v76* females to *nm38*/+ heterozygous males and also *nm38* females to *v76*/+ males. We then singled 50 L4 hermaphrodite F1 progeny away from their brothers from each cross to score their germlines under the dissecting microscope for the Fog phenotype; we also scored the germline phenotype of 50 F1 males from each mating with Nomarski optics for the presence of oocytes. If *nm38* and *trr*-1(v76) failed to complement each other, we would expect half of all XX and XO F1 animals to have a feminized germline. However, we instead found that all 100 XX worms were self-fertile and all 100 XO worms examined contained only sperm in their germlines. Thus, we conclude that *nm38* and *v76* do indeed complement each other and thus are not allelic.

To test *nm38* complementation of the *puf-2* deletion allele *nm66*, we mated 8 *cby-15* + / + *nm66* hermaphrodites from a strain created by Qinwen Liu (Haag laboratory) to 12 *nm38*

homozygous males at the larval stage L4 in bulk. Though the germline of *nm38* males eventually switches to oogenesis, after molting to adults they are capable of limited cross-fertility while still producing sperm. Plugged P_0 hermaphrodites were singled, and from plates that produced >35% males, we singled 50 L4 hermaphrodite F1 progeny to check their germline for the presence of self-sperm. Half of these singled F1 animals should be of the genotype + *puf-2*(*nm66*)/*nm38 +*, and if *nm38* and *nm66* fail to complement each other, these animals should have a feminized germline. However, we found that 50/50 XX F1s visually inspected were self-fertile. Thus, we conclude that *nm38* and *puf-2*(*nm66*) do indeed complement each other and are not allelic.

Mapping of nm38

To map the location of the *nm38* allele, we performed both phenotypic and molecular marker mapping. An initial *C. briggsae* genetic map has been created by Dr. Bhagwati Gupta (Gupta et al. 2007), and an extensive list of molecular mapping assays for *C. briggsae*, complete with computationally determined primers, was designed by R. Miller and D. Koboldt at Washington University School of Medicine in St. Louis, MO (Kobaldt 2010).

In phenotypic mapping, we used the mutations given in Table 3. (There were no phenotypic markers for Chr. V at the time of this work, and two additional markers on Chromosome 2, *cby-10*(sy5064) and *unc-4*(sy5341), proved not to be practical for phenotypic mapping.) Our results are given in Table 3. Scoring more than 1100 animals per assay, we found only linkage of *nm38* to chromosome 2.

Table 3

Chr.	Allele	wild-type animals	mutant-non Fogs	Fog non-mutants	double mutants	reject X_2	map distance
2	*dpy(nm4)*	940	312	311	15	Yes	19.5
2	*cby-15(sy5148)*	677	252	220	15	Yes	22.7
1	*lev-unc(sy5440)*	634	216	201	63	No	unlinked
3	*cby-4(s1272)*	919	320	282	97	No	unlinked
4	*cby-7(sy5027)*	706	275	242	82	No	unlinked

There are also two other phenotype mutations that map to chromosome 2, *bli*(sy5259) and *lin*(sy5342) with which we have not yet mapped. However, even taking these into account, there is only a low density of phenotypic markers suitable for further mapping on chromosome 2.

We also attempted to map *nm38* molecularly. We first used a bulk-segregant DNA mapping approach for indels and SNP-snips identified between AF16 and HK104 or VT847 (Koboldt et al. 2010). In this work, we examined bulk-segregant PCR products (restriction digested, if necessary) by eye for a large ratio of AF16 to HK104 or VT847 PCR products from F2 Fog worms in a mapping strain background, produced as above, compared to their WT siblings (who themselves should be slightly enriched for HK104/VT847 PCR products if the mapping assay location is close to *nm38*).

Initially using one or two assays along each chromosome, we found linkage of *nm38* to only chromosome 2. We then chose additional assays on chromosome 2 for finer mapping, including within fingerprint contigs (fpc's) that do not yet have a defined location on that chromosome (so-called "random" fpc's). As these assays are gel-based and thus not easily quantified, Table 4 provides a qualitative summary of this mapping work. We found the strongest linkage to chromosome 2 fpc 2260, with weaker linkage to the fpc's that flank it, fpc 1402 and 0305.

Table 4

Fingerprint contig assayed	Qualitative linkage score	Assay name and chromosome	Qualitative linkage score
Chromosome II		**Other chromosomes**	
fpc 0071a	none	gld-1(I)	None
fpc 2454a	weak	1.2 (I)	None
fpc 0058	weak/medium	atx-2 (III)	None
fpc 4206	medium	tra-1 (III)	None
fpc 1402	medium	3.1 (III)	None
fpc 2260	strongest	4.1 (IV)	None
fpc 0305	medium/strong	4.2 (IV)	None
random 4087	medium	5.1 (V)	None
random 4131	medium	X.1 (X)	None

Chromosome 2 fingerprint contigs are listed in their physical order, from 2MB to 14.2MB.

In order to quantify our molecular mapping results, we moved from bulk-segregant mapping to molecular mapping assays with individual animals. We performed these experiments using the mapping strain VT847, a *C. briggsae* strain more closely related to mutagenesis strain AF16 than HK104 (Cutter et al. 2006) to reduce the likelihood of marker non-colinearity between strains. First however, given both our inability to tightly link *nm38* to any marker in previous mapping attempts and the unusual genetic properties of the allele, we performed studies of the inheritance of the Fog phenotype in AF16 vs. VT847 backgrounds to account for genetic phenomenon like semi-dominance of *nm38* or modifiers within VT847 that would render this strain unsuitable for mapping. We mated *nm38* Fogs to either AF16 males as a control or to VT847 males. We scored 100 F1 hermaphrodite progeny from each kind of mating for general self-fertility and 20 F1 hermaphrodites for their ability to produce the expected number of *nm38* Fogs in the F2 generation. We found that 100% of animals assayed behaved as expected. Next, we singled 85 F2 XX L4 animals from the two different P_0 matings in order to

assess F3 phenotypes. We found 2/85 F2 animals to look Fog but produce small numbers of all Fog progeny (12 and 19 XX Fog animals, respectively) from the VT847 parental matings, and 1/85 F2 to look Fog but produce 8 Fog progeny from the AF16 parental matings. The other proportions of inferred F2 genotypes, *nm38*/*nm38*, *nm38*/+, and +/+, from the VT847 and AF16 parental matings are not statistically different from each other (X_2 = 2.1 respectively, 2 df). Additionally, all presumed *nm38* heterozygous mothers produced ~20% Fogs in the F3 generation. Thus we conclude that *nm38* behaves similarly in the AF16 and VT847 backgrounds.

Determining that there was no large-scale distortion in expected phenotypic ratios between AF16 and the mapping strain VT847, we proceeded with single-worm sequencing mapping assays. DNA was made from individual F2 Fog worms in the mapping strain background, as described above, and from their wild type-looking siblings for controls. We used this DNA as template in PCR and sequencing reactions and examined each sequence for the presence of AF16 or VT847 SNP alleles. We used mapping assays along chromosome 2 from fpc 4206 (at the 10MB position) to fpc 0305 at the end of the chromosome (14.2MB), which is the region to which we had seen the strongest *nm38* linkage by bulk-segregant analysis.

Figure 6 provides a summary of our individual Fog molecular mapping; arrows indicate the locations of some assays used, and the computed map distance (in map units) to *nm38* is given above each one.

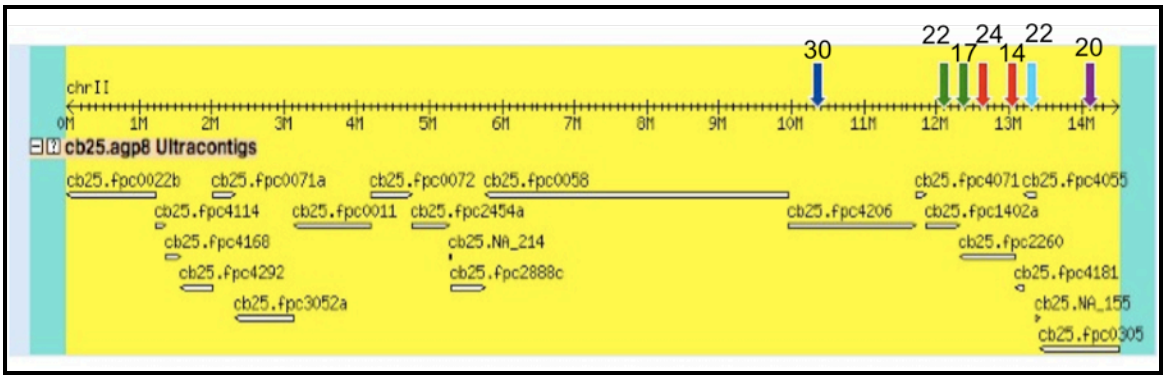


FIGURE 6. Fingerprint contig map of *C. briggsae* chromosome 2 from WormBase (build CB3). Colored arrows mark the locations of particular AF16-VT847 SNP mapping assays; identically colored arrows indicate assays predicted to lie on the same fpc. The numbers above each arrow are the number of map units of assay from *nm38*.

As can be seen, the strongest signal of linkage comes from the chromosomal region between 12.1 and 13.3MB. This is concordant with our bulk-segregant mapping data. Surprisingly, however, no one assay shows close linkage to *nm38*. More problematically, if the genetic versus physical distance in this region of chromosome 2 is really 10 centimorgans per megabase, as described by Hillier et al. (2007), then there is also not enough physical chromosome space in this genomic region to account for the genetic distances we find between markers. This observation is true even if we imagine some of the yet-unplaced, “random” fingerprint contigs (not shown) actually lie in this interval. A caveat to this work is that this end of chromosome 2 is assembled from a number of small fingerprint contigs (Figure 6): the rightmost three megabases of the chromosome contain seven fingerprint contigs, only two of which are larger than 800,000 bases. Thus, we are unable to more precisely determine the map location of *nm38*. As described below, this is perhaps either due to current genome mis-assembly or to genetic properties of *nm38* that render it un-mappable in these ways.

DISCUSSION

nm38 is a germline feminizing allele

We isolated one allele, *nm38*, in a forward genetic screen for mutations that feminize the *C. briggsae* hermaphrodite germline. The fate of sperm cells in XX and XO *nm38* animals is transformed into that of oocytes, though some presumed XX homozygotes have occasional limited self-fertility and XO homozygotes make some sperm before switching to oogenesis. As judged by immunohistochemistry, *nm38* XX Fogs do not make any detectable major sperm protein, despite the high sensitivity of the anti-MSP antibody used and the normal abundance of this protein (Kosinski et al. 2005). Although only noted, *nm38* likely has a maternal effect on embryonic/larval viability, and we present evidence that it has a maternally temperature sensitive effect on sex determination. *nm38* behaves in many ways like a recessive allele, though given the intersexual somas it can produce in certain crosses (Table 2), it could be an unusual antimorph, – a dominant negative mutation that is capable of poisoning wild-type copies of gene product.

nm38 can eliminate spermatogenesis in both Cbr-tra-2(lf) *and Cbr-fem-3*(lf) *mutants*

We analyzed double mutant phenotypes of *nm38* and two null or strong hypomorphic *C. briggsae* sex determination mutant alleles in order to elucidate genetic interactions between these sex determination loci. *tra-2* is a female-promoting gene, and loss-of-function alleles of *C. briggsae tra-2* result in masculinization of both the germline and soma of XX animals (Kelleher et al. 2008). In *C. elegans*, the genes *fem-1*, *fem-2*, and *fem-3* are male-promoting, necessary for male somatic and germline fates in XX and XO animals; in *C. briggsae,* however, though the somatic role of the *fem* genes is conserved, loss-of-function *fem* homozygous mutants retain hermaphrodite spermatogenesis (Hill *et al,* 2006). Constructing Fog(*nm38*) *Cbr-tra-2*(lf) and

Fog(*nm38*); *Cbr-fem-3*(*nm63*) animals, we find that each double mutant produces XX-karyotype animals with oocytes in their germlines and these animals look Fog. Thus, we conclude that *nm38* causes sperm to be transformed into oocytes even in the absence of the female-promoting *tra-2* gene product. Perhaps more unusual, *nm38* can also eliminate the residual masculinizing activity left in *fem-3(nm63)* XX animals, transforming those sperm to oocytes as well. As we do not know whether *nm38* is even a hypomorph, we cannot say anything about the true epistatic relationships between these alleles.

nm38 is not allelic to other known Fog mutations

nm38 is not allelic to orthologs of the *C. elegans* Fog genes *fog-1* or *fog-3*, nor is it allelic to the *C. briggsae* Fog genes *she-1*, *trr-1*, or *puf-2*. Therefore *nm38* might be an allele of a new sex determination gene, perhaps either a gene that is *C. briggsae*-specific, or a gene that is present in both *C. elegans* and *C. briggsae* but is only involved in *C. briggsae* sex determination. Alternatively, *nm38* might be an unusual allele, for instance with antimorphic properties, of a gene that can manifest a feminizing sex determination phenotype when altered this way. Lastly, given our difficulties mapping its physical location, *nm38* might be an unnatural genetic element created during mutagenesis, like a chromosome rearrangement or translocation that creates a Fog phenotype (see below).

nm38 shows unusual mapping behavior

Though we have localized *nm38* to the right arm of *C. briggsae* chromosome 2, we have not been able to obtain a finer map position than this by phenotypic, bulk-segregant, or individual molecular marker mapping assays. In particular, we find inflated map distances between various mapping markers and the *nm38* locus. There are potential explanations for

why *nm38* seems un-mappable in the ways we have undertaken. First, there could possibly be non-colinearity between our *nm38* strain and our mapping strain VT847 (and/or HK104) in the genomic region where *nm38* lies. That other *C. briggsae* forward mutations have been mapped on chromosome II, such as *Cbr-trr-1* (Dr. Ronald Ellis, personal communication) argues against wide-spread, background differences between AF16 and other *C. briggsae* strains on this chromosome. Perhaps, though, *nm38* is a chromosomal translocation of DNA from another chromosome onto chromosome 2. This would explain both its real linkage to chromosome 2 and also our inability to obtain markers on that chromosome that are closely linked to *nm38.* A formal check for co-linearity of phenotypic and molecular markers in this section of chromosome 2 should answer this question definitively.

Secondly, given the fragmented nature of its assembly in the genome sequence, it is possible that the region of chromosome 2 to which *nm38* shows the greatest linkage is misassembled or incompletely assembled. If *nm38* lies in DNA not properly accounted for on chromosome 2, then we would not be able to detect a strong signal of linkage to it with known physical markers.

Thirdly, it is also possible that *nm38* is a pleiotropic mutation causing not only the elimination of XX and XO spermatogenesis, but also a real increase of recombination rates, for instance by recruiting synaptonemal complex machinery to the chromosome or by encouraging double stranded breaks near the *nm38* locus. In this case, given that *C. elegans* and likely *C. briggsae* only have one successful crossover event per chromosome per meiosis, one would perhaps expect to see a decrease in recombination between other, weakly or un-linked chromosome II markers in an *nm38* mutant if recombination was instead forced to the vicinity of the *nm38* site. Alternatively, an *nm38*-mediated general increase in recombination might cause an increased recombination rate between all markers on chromosome II (Tsai et al. 2008,

Mets and Meyer 2009). We could test either of these two explanations by systematically determining map distances between phenotypic and/or molecular markers along chromosome 2 of the *nm38* strain wild-type *C. briggsae* to check for map distance inflation or deflation.

Fourthly, it is formally possible that despite extensive outcrossing, *nm38* is not a single mutational event (i.e., a single point mutation or small insertion/deletion mutation), but instead is a 'composite allele' made of two or more mutations that must be inherited together for viability. If such mutations are not tightly physically linked to one another other, a 'composite allele' may be expected to behave weirdly in mapping experiments. In this case, we would not expect *nm38* to show tight linkage to any one location in the *C. briggsae* genome.

A final possibility that would explain our consistently inflated mapping data is that *nm38* is not truly recessive, or not recessive specifically in a mapping strain genetic background. If this were true, then the Fog animals picked for each mapping experiment were not necessarily *nm38* homozygotes, but could have instead been *nm38* (AF16)/+ (mapping strain) Fog heterozygotes. However, in specific testing for the semi-dominance of *nm38* in any relevant genetic background, and for both dominant and recessive enhancers of the Fog phenotype in HK104 and VT847, we found no strong evidence for this hypothesis.

Future work to be done with nm38

With the creation of a robust *nm38* homozygous stock, it is now possible to conduct straightforward genetic screens for phenotypic mutations (dpys, uncs, etc.) linked to *nm38*. With such a marked *nm38* strain, it might be feasible to isolate other alleles of the *nm38* locus in non-complementation screens. Non-complementation screens are an important part of genetic analysis in that they allow fuller characterization of a locus through the study of different mutations at that locus that may have more or less severe, or qualitatively different, properties.

Further, they allow one to identify characteristics of an allele that are allele-specific versus ones that are truly of that locus. We considered making non-complementation screens a priority in our *nm38* work, but ultimately decided that the mapping difficulties made it unacceptably likely that *nm38* was not a standard allele amenable to this approach.

A robust *nm38* homozygous stock also makes it possible to perform suppressor screens for mutations that restore self-fertility to *nm38* Fogs. Again, as we are wary of the un-mappable nature of *nm38*, and also because, unfortunately, *C. briggsae* still has relatively few phenotypic markers to aid in the mapping of suppressors, we thought this experiment should also not be a top priority. But, as for non-complementation screens described above, a suppressor screen should certainly be considered in the future once the allelic nature of *nm38* becomes clearer and genetic mapping in *C. briggsae* becomes more feasible.

APPENDIX 1:

Injections were into the guts of adult AF16 hermaphrodites grown at 20°C and recovered to 20°C.
Adult F1 progeny were scored using a dissecting microscope and with DIC microscopy if necessary.
Unless noted, the penetrance of phenotypes ranged between ~10% and 33%.
Entries with no description produced no mutant phenotypes.

Gene 1	Gene 2	Obvious mutant phenotypes
CBG02483	CBG09734	Protruding vulva. Many fewer germ cells than WT. Germ cell arrest. Very few gametes those that form (sperm/oocytes) look malformed. Very small /no germline mitotic zone
Cbr-pie-1	CBG03777	Slow ovulation/fertilization/embryo laying.
Cbr-spn-4	*Cbr-set-14*	
Cbr-skr-1	*Cbr-hop-1*	
Cbr-pos-1	CBG07193	
CBG09898	*Cbr-daz-1*	Slow ovulation/fertilization/embryo laying.
Cbr-pal-1	*Cbr-rme-2*	Aberrant oocytes.
Cbr-oma	*Cbr-tag-246*	Unusual uterine tissue appearance ('fragile, sloppy'). Some large but misshapen and weirdly granular oocytes. Few embryos in uteri, maybe due to slow ovulation.
Cbr-arl-8	*Cbr-swd-3.3*	
Cbr-patr-1	*Cbr-rskn-1*	
Cbr-puf-"4"	*Cbr-puf-8*	Aberrantly small, misshapen spermatozoa.
CBG22317	*Cbr-nos-2*	
CBG08921	CBG04207	Slow growth.
CBG05292	CBG01393	
CBG16726	CBG05879	
CBG08527	CBG11569	
CBG05095	CBG09653	
CBG10091	CBG13227	
CBG04372	CBG04373	
CBG03080	CBG09925	
CBG20875	CBG14962	
Cbr-egg-4	*Cbr-tpa-1*	Injected worms lay only oocytes.
CBG00199	CBG22683	
CBG01956	CBG02251	Slow ovulation/fertilization/embryo laying.
CBG0282R	CBG03076	Aberrant oocytes, ooids. 'Disintegrating' proximal oocytes.
CBG03085	CBG03615	No germ cells.
CBG04301	CBG04302	Oocytes are mis-formed or are 'missing' (just gonad basement membrane is visible).
CBG21596	CBG20384	Oocytes are mis-formed or are 'missing' (just gonad basement membrane is visible).
CBG14085	*Cbr-aly-1*	Aberrant oocytes.
CBG13508	CBG07045	Aberrant oocytes.
CBG07640	CBG07661	Aberrant oocytes, and 'disintegration' of the most proximal oocyte.
CBG08571	CBG08989	Aberrant oocytes, with ooids.
CBG09062	CBG09108	Delayed gametogenesis': sperm and residual bodies in adults, with only ~2 grainy oocytes/gonad arm. Normal pachytene region, transition zone and mitotic region.
CBG09250	CBG09264	Aberrant oocytes.
CBG09348	CBG09840	Many defects: few germ cells, aberrant oocytes, few sperm in gonad arms but non-stacking oocytes.
CBG10477	CBG10809	Aberrant proximal oocytes, and post-pachytene cells also look abnormal.
CBG11199	*Cbr-mes-3*	

CBG20654	*Cbr-moe-3*	Slow growth. Many defects: small and egg-laying defective adults, abnormal gonad migration and somatic gonad, protruding vulva, aberrant oogenesis.
CBG12306	CBG07050	A little gametogenesis delay, aberrant oocytes, 'disintegrating' proximal oocyte, initial embryonic cell divisions are abnormal (cleavage planes and location of nuclei).
Cbr-puf-"11"	*Cbr-puf-"4"*	
CBG13131	*Cbr-mop-25.3*	Aberrant oocytes. Laying oocytes (without oocyte stacking).
Cbr-lir-1	CBG03256	Slow ovulation/fertilization/embryo laying. Low penetrance oocyte defects.
CBG06213	*Cbr-mex-3*	100% embryonic lethality.
Cbr-glp-1	*Cbr-nhr-43*	100% embryonic lethality.
CBG05635	CBG05978	
CBG02511	CBG02683	Aberrant oocytes. Laying oocytes (without oocyte stacking).
Cbr-dmd-6	*Cbr-unc-71*	
CBG01946	CBG09113	Low penetrance germ cell decay through whole germline.
CBG04364	CBG11013	Decaying distal and proximal germline, aberrant oocytes.
CBG05916	CBG11273	Aberrant oocytes and some 'disintegrating' oocytes.
Cbr-alg-2		

Single RNAi injections

Cbr-arl-8	Nearly 100% embryonic lethality. Living adult progeny have aberrant oocytes.
CBG22317	Aberrant oocytes.
Cbr-nos-2	Aberrant oocytes, sometimes abnormal somatic gonad migration, rarely no germline.
CBG04207	Mildly aberrant oocytes, with a weird texture.
CBG08921	Aberrant oocytes.
Cbr-swd-3.3	Aberrant oocytes.
CBG10091	Decaying proximal oocytes.
CBG13227	Mildly aberrant and disorganized oocytes. Laying oocytes without stacking.
Cbr-tpa-1	Aberrant oocytes.
CBG03085	Nearly 100% penetrant aberrant proximal oocytes and post-pachytene cells.
CBG03615	Aberrant proximal oocytes and post-pachytene cells.
CBG13131	Slightly aberrantly shaped oocytes.
Cbr-mop-25.3	No germline mutant phenotype.
Cbr-nhr-43	No germline mutant phenotype.
CBG02511	No germline mutant phenotype.
CBG02683	Very weakly penetrant aberrant oogenesis.

REFERENCES:

Aguinaldo AM, Turbeville JM, Linford LS, Rivera MC, Garey JR, Raff RA, Lake JA (1997) Evidence for a clade of nematodes, arthropods and other moulting animals. Nature 387(6632):489-93.

Arning S, Groter P, Bilbe G, Kramer A Mammalian splicing factor SF1 is encoded by variant cDNAs and binds to RNA. RNA 2(8):794-810.

Baehrecke EH (1997) who encodes a KH RNA binding protein that functions in muscle development. Development 124(7):1323-32.

Barske LA, Capel B (2008) Blurring the edges in vertebrate sex determination. Curr Opin Genet Dev 18(6):499-505.

Barton MK, Kimble J (1990) *fog-1*, a regulatory gene required for specification of spermatogenesis in the germ line of *Caenorhabditis elegans*. Genetics 125(1):29-39.

Beuck C, Szymczyna BR, Kerkow DE, Carmel AB, Columbus L, Stanfield RL, Williamson JR (2010) Structure of the GLD-1 homodimerization domain: insights into STAR protein-mediated translational regulation. Structure 18(3):377-89.

Bull, J (1983) The genome sequence of *Caenorhabditis* briggsae: a platform for comparative genomics. Evolution of sex determining mechanisms. Menlo Park, CA,
Benjamin Cummings Publishing Co.

Cali BM, Anderson P (1998) mRNA surveillance mitigates genetic dominance in *Caenorhabditis elegans*. Mol Gen Genet 260:176-84.

Chawla G, Lin CH, Han A, Shiue L, Ares M Jr, Black DL (2009) Sam68 regulates a set of alternatively spliced exons during neurogenesis. Mol Cell Biol 29(1):201-13.

Chen PJ, Cho S, Jin SW, Ellis RE (2001) Specification of germ cell fates by FOG-3 has been conserved during nematode evolution. Genetics 158(4):1513-25.

Chen PJ, Singal A, Kimble J, Ellis RE (2000) A novel member of the tob family of proteins controls sexual fate in *Caenorhabditis elegans* germ cells. Dev Biol 217(1):77-90.

Chen T, Damaj BB, Herrera C, Lasko P, Richard S (1997) Self-association of the single-KH-domain family members Sam68, GRP33, GLD-1, and Qk1: role of the KH domain. Mol Cell Biol 17(10):5707-18.

Chen T, Richard S (1998) Structure-function analysis of Qk1: a lethal point mutation in mouse quaking prevents homodimerization. Mol Cell Biol 18(8):4863-71.

Chin-Sang ID, Spence AM (1996) *Caenorhabditis elegans* sex-determining protein FEM-2 is a protein phosphatase that promotes male development and interacts directly with FEM-3. Genes Dev 10(18):2314-25.

Cline TW and Meyer BJ (1996) Vive la difference: males vs females in flies vs worms. Annu Rev Genet 30: 637-702.

Collins L, Penny D (2005) Complex spliceosomal organization ancestral to extant eukaryotes. Mol Biol Evol 22(4):1053-66.

Cutter AD (2008) Divergence times in *Caenorhabditis* and *Drosophila* inferred from direct estimates of the neutral mutation rate. Mol Biol Evol 25:778-786.

Cutter AD, Felix MA, Barriere A, Charlesworth D (2006) Patterns of nucleotide polymorphism distinguish temperate and tropical wild isolates *of Caenorhabditis briggsae*. Genetics 173(4):2021-31.

Dabney AR, Storey JD (2007) A new approach to intensity-dependent normalization of two-channel microarrays. Biostatistics 8(1):128-39.

Dabney AR, Storey JD (2007) Normalization of two-channel microarrays accounting for experimental design and intensity-dependent relationships. Genome Biol 8(3):R44

de Bono M, Hodgkin J (1996) Evolution of sex determination in *Caenorhabditis*: unusually high divergence of *tra-1* and its functional consequences. Genetics 144(2):587-95.

Di Fruscio M, Chen T, Bonyadi S, Lasko P, Richard S (1998) The identification of two Drosophila K homology domain proteins. Kep1 and SAM are members of the Sam68 family of GSG domain proteins. J Biol Chem 273(46):30122-30.

Di Fruscio M, Styhler S, Wikholm E, Boulanger MC, Lasko P, Richard S (2003) Kep1 interacts genetically with dredd/caspase-8, and kep1 mutants alter the balance of dredd isoforms. Proc Natl Acad Sci U S A 100(4):1814-9.

Doniach T (1986) Activity of the sex-determining gene *tra-2* is modulated to allow spermatogenesis in the *C. elegans* hermaphrodite. Genetics 114(1):53-76.

Dopazo H, Dopazo J (2005) Genome-scale evidence of the nematode-arthropod clade. Genome Biol 6(5):R41.

Ebersole TA, Chen Q, Justice MJ, Artzt K (1996) The quaking gene product necessary in embryogenesis and myelination combines features of RNA binding and signal transduction proteins. Nat Genet 12(3):260-5.

Ellis RE (2008) Sex determination in the *Caenorhabditis elegans* germ line. Curr Top Dev Biol 83:41-64.

Ellis RE, Kimble J (1995) The *fog-3* gene and regulation of cell fate in the germ line of *Caenorhabditis elegans*. Genetics 139(2):561-77.

Fay D, ed. (2006) Genetic mapping and manipulation, WormBook, doi/10.1895/wormbook.1.90.1, http://www.wormbook.org.

Francis R, Barton MK, Kimble J, Schedl T (1995) *gld-1*, a tumor suppressor gene required for oocyte development in *Caenorhabditis elegans*. Genetics 139(2):579-606.

Francis R, Maine E, Schedl T (1995) Analysis of the multiple roles of *gld-1* in germline development: interactions with the sex determination cascade and the glp-1 signaling pathway. Genetics 139(2):607-30.

Galarneau A, Richard S (2005) Target RNA motif and target mRNAs of the Quaking STAR protein. Nat Struct Mol Biol 12(8):691-8.

Galarneau A, Richard S (2009) The STAR RNA binding proteins GLD-1, QKI, SAM68 and SLM-2 bind bipartite RNA motifs. BMC Mol Biol 10:47.

Gonczy P, Echeverri C, Oegema K, Coulson A, Jones SJ, Copley RR, Duperon J, Oegema J, Brehm M, Cassin E, Hannak E, Kirkham M, Pichler S, Flohrs K, Goessen A, Leidel S, Alleaume AM, Martin C, Ozlu N, Bork P, Hyman AA (2000) Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. Nature 408(6810):331-6.

Goodwin EB, Okkema PG, Evans TC, Kimble J (1993) Translational regulation of *tra-2* by its 3' untranslated region controls sexual identity in *C. elegans.* Cell 75(2):329-39.

Grant B, Hirsh D (1999) Receptor-mediated endocytosis in the *Caenorhabditis elegans* oocyte. Mol Biol Cell 10(12):4311-26.

Guo Y, Lang S, Ellis RE (2009) Independent recruitment of F box genes to regulate hermaphrodite development during nematode evolution. Curr Biol 19(21):1853-60.

Gupta BP, et al. (2007) Genomics and biology of the nematode *Caenorhabditis briggsae*, WormBook, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.136.1, http://www.wormbook.org.

Haag ES (2005) The evolution of nematode sex determination: *C. elegans* as a reference point for comparative biology, WormBook, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.120.1, http://www.wormbook.org.

Haag ES, Kimble J (2000) Regulatory elements required for development of *Caenorhabditis elegans* hermaphrodites are conserved in the *tra-2* homologue of *C. remanei,* a male/female sister species. Genetics 155:1485.

Hansen D, Schedl T (2006) The regulatory network controlling the proliferation-meiotic entry decision in the *Caenorhabditis elegans* germ line. Curr Top Dev Biol 76:185-215.

Hardy RJ, Loushin CL, Friedrich VL Jr, Chen Q, Ebersole TA, Lazzarini RA, Artzt K (1996) Neural cell type-specific expression of QKI proteins is altered in quakingviable mutant mice. J Neurosci 16(24):7941-9.

Hill RC, de Carvalho CE, Salogiannis J, Schlager B, Pilgrim D, Haag ES (2006) Genetic flexibility in convergent evolution of hermaphroditism in *Caenorhabditis* nematodes. Dev Cell 10(4): 531-8.

Hillier LW, Miller RD, Baird SE, Chinwalla A, Fulton LA, Koboldt DC, Waterston RH (2007) Comparison of *C. elegans* and *C. briggsae* genome sequences reveals extensive conservation of chromosome organization and synteny. PLoS Biol 5(7):e167.

Hodgkin J (1986) Sex determination in the nematode *C. elegans*: analysis of *tra-3* suppressors and characterization of fem genes. Genetics 114(1):15-52.

Hodgkin JA, Brenner S (1977) Mutations causing transformation of sexual phenotype in the nematode *Caenorhabditis elegans*. Genetics 86(2):275-87.

Huelsenbeck JP and Ronquist F (2001) MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17:754-755.

Jan E, Motzny CK, Graves LE, Goodwin EB (1999) The STAR protein, GLD-1, is a translational regulator of sexual identity in *Caenorhabditis elegans*. EMBO J 18(1):258-69.

Janes DE, Organ CL, Edwards SV (2010) Variability in sex-determining mechanisms influences genome complexity in Reptilia. Cytogenet Genome Res [Epub ahead of print]

Janzen FJ, Phillips PC (2006) Exploring the evolution of environmental sex determination, especially in reptiles. J Evol Biol 19(6):1775-84.

Jarne P, Auld JR (2006) Animals mix it up too: the distribution of self-fertilization among hermaphroditic animals. Evolution 60(9):1816-24.

Jin SW, Kimble J, Ellis RE (2001) Regulation of cell fate in *Caenorhabditis elegans* by a novel cytoplasmic polyadenylation element binding protein. Dev Biol 229(2):537-53.

Jones AR, Francis R, Schedl T (1996) GLD-1, a cytoplasmic protein essential for oocyte differentiation, shows stage- and sex-specific expression during *Caenorhabditis elegans* germline development. Dev Biol 180(1):165-83.

Jones AR, Schedl T (1995) Mutations in *gld-1*, a female germ cell-specific tumor suppressor gene in *Caenorhabditis elegans*, affect a conserved domain also found in Src-associated protein Sam68. Genes Dev 9(12):1491-504.

Kelleher DF, de Carvalho CE, Doty AV, Layton M, Cheng AT, Mathies LD, Pilgrim D, Haag ES (2008) Comparative genetics of sex determination: masculinizing mutations in *Caenorhabditis briggsae*. Genetics 178(3):1415-29.

Kelly WG, Schaner CE, Dernburg AF, Lee MH, Kim SK, Villeneuve AM, Reinke V (2002) X-chromosome silencing in the germline of *C. elegans.* Development 129(2):479-92.

Kershner AM, Kimble J (2010) Genome-wide analysis of mRNA targets for *Caenorhabditis elegans* FBF, a conserved stem cell regulator. Proc Natl Acad Sci U S A 107:3936-3941.
Killian DJ, Hubbard EJ (2005) *Caenorhabditis elegans* germline patterning requires coordinated development of the somatic gonadal sheath and the germ line. Dev Biol 279(2):322-35.

Kim KW, Nykamp K, Suh N, Bachorik JL, Wang L, Kimble J (2009) Antagonism between GLD-2 binding partners controls gamete sex. Dev Cell 16(5):723-33.

Kimble J, Crittenden SL (2007) Controls of germline stem cells, entry into meiosis, and the sperm/oocyte decision in *Caenorhabditis elegans*. Annu Rev Cell Dev Biol 23:405-33.

Kimble J, Edgar L, Hirsh D (1984) Specification of male development in *Caenorhabditis elegans*: the fem genes. Dev Biol 105(1):234-9.

King N, Westbrook MJ, Young SL, Kuo A, Abedin M, Chapman J, Fairclough S, Hellsten U, Isogai Y, Letunic I, Marr M, Pincus D, Putnam N, Rokas A, Wright KJ, Zuzow R, Dirks W, Good M, Goodstein D, Lemons D, Li W, Lyons JB, Morris A, Nichols S, Richter DJ, Salamov A, Sequencing JG, Bork P, Lim WA, Manning G, Miller WT, McGinnis W, Shapiro H, Tjian R, Grigoriev IV, Rokhsar D (2008) The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. Nature 451(7180):783-788.

Kiontke K and Fitch DHA (2005) The Phylogenetic relationships of *Caenorhabditis* and other rhabditids, WormBook, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.11.1, http://www.wormbook.org.

Kiontke K and Sudhaus W (2006) Ecology of *Caenorhabditis* species, WormBook, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.37.1, http://www.wormbook.org.

Kiontke K, Gavin NP, Raynes Y, Roehrig C, Piano F, Fitch DH (2004) *Caenorhabditis* phylogeny predicts convergence of hermaphroditism and extensive intron loss. Proc Natl Acad Sci U S A 101(24):9003-8.

Koboldt DC, Staisch J, Thillainathan B, Haines K, Baird SE, Chamberlin HM, Haag ES, Miller RD, Gupta BP (2010) A toolkit for rapid gene mapping in the nematode Caenorhabditis briggsae. BMC Genomics. 2010 Apr 13;11:236.

Kosinski M, McDonald K, Schwartz J, Yamamoto I, Greenstein D (2005) *C. elegans* sperm bud vesicles to deliver a meiotic maturation signal to distant oocytes. Development 132(15):3357-69.

Kramer A, Utans U (1991) Three protein factors (SF1, SF3 and U2AF) function in pre-splicing complex formation in addition to snRNPs. EMBO J 10(6):1503-9.

Kuwabara PE, Kimble J (1995) A predicted membrane protein, TRA-2A, directs hermaphrodite development in *Caenorhabditis elegans*. Development 121(9):2995-3004.

L'Hernault SW Spermatogenesis (2006) WormBook, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.85.1, http://www.wormbook.org.

Lamont LB, Crittenden SL, Bernstein D, Wickens M, Kimble J (2004) FBF-1 and FBF-2 regulate the size of the mitotic region in the *C. elegans* germline. Dev Cell 7(5):697-707.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007). Clustal W and Clustal X version 2.0. Bioinform 23, 2947-2948.

Larocque D, Pilotte J, Chen T, Cloutier F, Massie B, Pedraza L, Couture R, Lasko P, Almazan G, Richard S (2002) Nuclear retention of MBP mRNAs in quaking viable mice. Neuron 36(5):815-29.

Larocque D, Richard S (2005) QUAKING KH domain proteins as regulators of glial cell fate and myelination. RNA Biol 2(2):37-40.

Lee MH and Schedl T. (2006) RNA-binding proteins, WormBook, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.79.1, http://www.wormbook.org.

Lee MH, Schedl T (2001) Identification of in vivo mRNA targets of GLD-1, a maxi-KH motif containing protein required for *C. elegans* germ cell development. Genes Dev 15(18):2408-20.

Lee MH, Schedl T (2004) Translation repression by GLD-1 protects its mRNA targets from nonsense-mediated mRNA decay in *C. elegans*. Genes Dev 18(9):1047-59.

Leek JT, Monsen E, Dabney AR, Storey JD (2006) EDGE: extraction and analysis of differential gene expression. Bioinformatics 22(4):507-508.

Lehmann-Blount KA, Williamson JR (2005) Shape-specific nucleotide binding of single-stranded RNA by the GLD-1 STAR domain. J Mol Biol 346(1):91-104.

Li J, Liu Y, Kim BO, He JJ (2002) Direct participation of Sam68, the 68-kilodalton Src-associated protein in mitosis, in the CRM1-mediated Rev nuclear export pathway. J Virol 76(16):8374-82.

Li Z, Zhang Y, Li D, Feng Y (2000) Destabilization and mislocalization of myelin basic protein mRNAs in quaking dysmyelination lacking QKI RNA-binding proteins. J Neurosci 20(13):4944-53.

Liu Z, Luyten I, Bottomley MJ, Messias AC, Houngninou-Molango S, Sprangers R, Zanier K, Kramer A, Sattler M (2001) Structural basis for recognition of the intron branch site RNA by splicing factor 1. Science 294(5544):1098-102.

Lo PC, Frasch M (1997) A novel KH-domain protein mediates cell adhesion processes in Drosophila. Dev Biol 190(2):241-56.

Lublin AL, Evans TC (2007) The RNA-binding proteins PUF-5, PUF-6, and PUF-7 reveal multiple systems for maternal mRNA regulation during *C. elegans* oogenesis. Dev Biol 303(2):635-49.

Lukong KE, Richard S (2003) Sam68, the KH domain-containing superSTAR. Biochim Biophys Acta 1653(2):73-86.

Lukong KE, Richard S (2008) Motor coordination defects in mice deficient for the Sam68 RNA-binding protein. Behav Brain Res 189(2):357-63.

Maguire ML, Guler-Gane G, Nietlispach D, Raine AR, Zorn AM, Standart N, Broadhurst RW (2005) Solution structure and backbone dynamics of the KH-QUA2 region of the Xenopus STAR/GSG quaking protein. J Mol Biol 348(2):265-79.

Maine EM, Hansen D, Springer D, Vought VE (2004) *Caenorhabditis elegans atx-2* promotes germline proliferation and the oocyte fate. Genetics 168(2):817-30.

Mank JE, Avise JC (2009) Evolutionary diversity and turn-over of sex determination in teleost fishes. Sex Dev 3(2-3):60-7.

Marin VA, Evans TC (2003) Translational repression of a *C. elegans* Notch mRNA by the STAR/KH domain protein GLD-1. Development 130(12):2623-32.

Marshall Graves JA (2008) Weird animal genomes and the evolution of vertebrate sex and sex chromosomes. Annu Rev Genet 42:565-86.

Matter N, Herrlich P, Konig H (2002) Signal-dependent regulation of splicing via phosphorylation of Sam68. Nature 420(6916):691-5.

McCarter J, Bartlett B, Dang T, and Schedl T (1999) On the control of oocyte meiotic maturation and ovulation in *Caenorhabditis elegans*. Dev Biol 205: 111-128.

Mehra A, Gaudet J, Heck L, Kuwabara PE, Spence AM (1999) Negative regulation of male development in *Caenorhabditis elegans* by a protein-protein interaction between TRA-2A and FEM-3. Genes Dev 13(11):1453-63.

Meiklejohn CD, Tao Y (2010) Genetic conflict and sex chromosome evolution. Trends Ecol Evol 25(4):215-223.

Mets DG, Meyer BJ (2009) Condensins regulate meiotic DNA break distribution, thus crossover frequency, by controlling chromosome structure. Cell 139:73-86.

Mootz D, Ho DM, Hunter CP (2004) The STAR/Maxi-KH domain protein GLD-1 mediates a developmental switch in the translational control of *C. elegans* PAL-1. Development 131:3263-3272.

Nabel-Rosen H, Volohonsky G, Reuveny A, Zaidel-Bar R, Volk T (2002) Two isoforms of the Drosophila RNA binding protein, how, act in opposing directions to regulate tendon cell differentiation. Dev Cell 2(2):183-93.

Najib S, Martin-Romero C, Gonzalez-Yanes C, Sanchez-Margalet V (2005) Role of Sam68 as an adaptor protein in signal transduction. Cell Mol Life Sci 62(1):36-43.

Nayak S, Goree J, Schedl T (2005) *fog-2* and the evolution of self-fertile hermaphroditism in *Caenorhabditis*. PLoS Biol 3(1):e6.

Noble SL, Allen BL, Goh LK, Nordick K, Evans TC (2008) Maternal mRNAs are regulated by diverse P body-related mRNP granules during early *Caenorhabditis elegans* development. J Cell Biol 182(3):559-72.

Ohno G, Hagiwara M, Kuroyanagi H (2008) STAR family RNA-binding protein ASD-2 regulates developmental switching of mutually exclusive alternative splicing in vivo. Genes Dev 22:360-374.

Okkema PG, Kimble J (1991) Molecular analysis of *tra-2*, a sex determining gene in *C. elegans*. EMBO J 10(1):171-6.

Oliveros JC (2007) VENNY. An interactive tool for comparing lists with Venn Diagrams. http://bioinfogp.cnb.csic.es/tools/venny/index.html.

Paronetto MP, Bianchi E, Geremia R, Sette C (2008) Dynamic expression of the RNA-binding protein Sam68 during mouse pre-implantation development. Gene Expr Patterns 8(5):311-22.

Paronetto MP, Messina V, Bianchi E, Barchi M, Vogel G, Moretti C, Palombi F, Stefanini M, Geremia R, Richard S, Sette C (2009) Sam68 regulates translation of target mRNAs in male germ cells, necessary for mouse spermatogenesis. J Cell Biol 185(2):235-49.

Paronetto MP, Zalfa F, Botti F, Geremia R, Bagni C, Sette C (2006) The nuclear RNA-binding protein Sam68 translocates to the cytoplasm and associates with the polysomes in mouse spermatocytes. Mol Biol Cell 17(1):14-24.

Peterson KJ, Eernisse DJ (2001) Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18S rDNA gene sequences. Evol Dev. May-Jun;3(3):170-205.

Philippe H, Lartillot N, Brinkmann H (2005) Multigene analyses of bilaterian animals corroborate the monophyly of Ecdysozoa, Lophotrochozoa, and Protostomia. Mol Biol Evol 22(5):1246-53.

Pilgrim D, McGregor A, Jackle P, Johnson T, Hansen D (1995) The *C. elegans* sex-determining gene *fem-2* encodes a putative protein phosphatase. Mol Biol Cell 6(9):1159-71.

Pilotte J, Larocque D, Richard S (2001) Nuclear translocation controlled by alternatively spliced isoforms inactivates the QUAKING apoptotic inducer. Genes Dev 15(7):845-58.

Pires-daSilva A (2007) Evolution of the control of sexual identity in nematodes. Semin Cell Dev Biol 18(3):362-70.

Praitis V, Casey E, Collar D, Austin J (2001) Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. Genetics 157(3):1217-26.

Ramakers C, Ruijter JM, Deprez RH, Moorman AF (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci Lett 339(1):62-6.

Reddy TR (2000) A single point mutation in the nuclear localization domain of Sam68 blocks the Rev/RRE-mediated transactivation. Oncogene 19(27):3110-4.

Reinke V, Gil IS, Ward S, Kazmer K (2004) Genome-wide germline-enriched and sex-biased expression profiles in *Caenorhabditis elegans*. Development 131(2):311-23.

Ritchie ME, Silver J, Oshlack A, Holmes M, Diyagama D, Holloway A, Smyth GK (2007) A comparison of background correction methods for two-colour microarrays. Bioinformatics 23(20):2700-7.

Robard C, Daviau A, Di Fruscio M (2006) Phosphorylation status of the Kep1 protein alters its affinity for its protein binding partner alternative splicing factor ASF/SF2. Biochem J 400(1):91-7.

Ronquist F and Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

Rosenquist TA, Kimble J (1988) Molecular cloning and transcript analysis of *fem-3,* a sex-determination gene in *Caenorhabditis elegans*. Genes Dev 2(5):606-16.

Ruijter JM, Ramakers C, Hoogaars WM, Karlen Y, Bakker O, van den Hoff MJ, Moorman AF (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Res 37(6):e45.

Russell S, Meadows L, Russell R, ed. (2009) Microarray Technology in Practice. Elsevier Burlington, MA.

Ryder SP, Frater LA, Abramovitz DL, Goodwin EB, Williamson JR (2004) RNA target specificity of the STAR/GSG domain post-transcriptional regulatory protein GLD-1. Nat Struct Mol Biol 11(1):20-8.

Ryder SP, Williamson JR (2004) Specificity of the STAR/GSG domain protein Qk1: implications for the regulation of myelination. RNA 10(9):1449-58.

Saccomanno L, Loushin C, Jan E, Punkay E, Artzt K, Goodwin EB (1999) The STAR protein QKI-6 is a translational repressor. Proc Natl Acad Sci U S A 96(22):12605-10.

SC Weeks, C Benvenuto, and SK Reed (2006) When males and hermaphrodites coexist: a review of androdioecy in animals. Integrative and Comparative Biology 46:449-464.

Schedl T, Kimble J (1988) *fog-2,* a germ-line-specific sex determination gene required for hermaphrodite spermatogenesis in *Caenorhabditis elegans*. Genetics 119(1):43-61.

Selenko P, Gregorovic G, Sprangers R, Stier G, Rhani Z, Kramer A, Sattler M (2003) Structural basis for the molecular recognition between human splicing factors U2AF65 and SF1/mBBP. Mol Cell 11(4):965-76.

Silver JD, Ritchie ME, Smyth GK (2009) Microarray background correction: maximum likelihood estimation for the normal-exponential convolution. Biostatistics 10(2):352-363.

Sinclair AH, Smith C et al. (2002) A comparative analysis of vertebrate sex determination, from The Genetics and Biology of Sex Determination (Novartis Found. Symp. 244). Chichester, UK, Wiley Publishing.

Stein LD, Bao Z, Blasiar D, Blumenthal T, Brent MR, Chen N, Chinwalla A, Clarke L, Clee C, Coghlan A, Coulson A, D'Eustachio P, Fitch DH, Fulton LA, Fulton RE, Griffiths-Jones S, Harris TW, Hillier LW, Kamath R, Kuwabara PE, Mardis ER, Marra MA, Miner TL, Minx P, Mullikin JC, Plumb RW, Rogers J, Schein JE, Sohrmann M, Spieth J, Stajich JE, Wei C, Willey D, Wilson RK, Durbin R, Waterston RH (2003) The genome sequence of *Caenorhabditis* briggsae: a platform for comparative genomics. PLoS Biol 1(2):E45.

Storey JD, Dai JY, Leek JT (2007) The optimal discovery procedure for large-scale significance testing, with applications to comparative microarray experiments. Biostatistics 8(2):414-32.

Streit A, Li W, Robertson B, Schein J, Kamal IH, Marra M, Wood WB (1999) Homologs of the *Caenorhabditis elegans* masculinizing gene *her-1* in C. briggsae and the filarial parasite Brugia malayi. Genetics 152(4):1573-84.

Subramaniam K, Seydoux G (2003) Dedifferentiation of primary spermatocytes into germ cell tumors in *C. elegans* lacking the pumilio-like protein PUF-8. Curr Biol 13(2):134-9.

Taylor SJ, Resnick RJ, Shalloway D (2004) Sam68 exerts separable effects on cell cycle progression and apoptosis. BMC Cell Biol 5:5.

Taylor SJ, Shalloway D (1994) An RNA-binding protein associated with Src through its SH2 and SH3 domains in mitosis. Nature 368(6474):867-71.

Tenenbaum SA, Carson CC, Lager PJ, Keene JD (2000) Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays. Proc Natl Acad Sci U S A 97(26):14085-90.

The Gene Ontology Consortium (2000) Gene Ontology: tool for the unification of biology. Nature Genet 25:25-29.

Thompson SR, Goodwin EB, Wickens M (2000) Rapid deadenylation and Poly(A)-dependent translational repression mediated by the *Caenorhabditis elegans* tra-2 3' untranslated region in Xenopus embryos. Mol Cell Biol 20(6):2129-37.

Tremblay GA, Richard S (2006) mRNAs associated with the Sam68 RNA binding protein. RNA Biol 3(2):90-3.

Tsai CJ, Mets DG, Albrecht MR, Nix P, Chan A, Meyer BJ (2008) Meiotic crossover number and distribution are regulated by a dosage compensation protein that resembles a condensin subunit. Genes Dev 22:194-211.

Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 98(9):5116-5121.

Valenzuela N (2008) Sexual development and the evolution of sex determination. Sex Dev 2(2):64-72.

Vernet C, Artzt K (1997) STAR, a gene family involved in signal transduction and activation of RNA. Trends Genet 13(12):479-84.

Volk T, Israeli D, Nir R, Toledano-Katchalski H (2008) Tissue development and RNA control: "HOW" is it coordinated? Trends Genet 24(2):94-101.

Wallis MC, Waters PD, Graves JA (2008) Sex determination in mammals--before and after the evolution of SRY. Cell Mol Life Sci 65(20):3182-95.

Ward S, Roberts TM, Strome S, Pavalko FM, Hogan E (1986) Monoclonal antibodies that recognize a polypeptide antigenic determinant shared by multiple *Caenorhabditis elegans* sperm-specific proteins. J Cell Biol 102(5):1778-86.

Wettenhall JM, Smyth GK (2004) limmaGUI: a graphical user interface for linear modeling of microarray data. Bioinformatics 20(18): 3705-3706.

Wu JI, Reed RB, Grabowski PJ, Artzt K (2002) Function of quaking in myelination: regulation of alternative splicing. Proc Natl Acad Sci U S A 99(7):4233-8.

Xu L, Paulsen J, Yoo Y, Goodwin EB, Strome S (2001) *Caenorhabditis elegans* MES-3 is a target of GLD-1 and functions epigenetically in germline development. Genetics 159:1007-1017.

Yanai I, Hunter CP (2009) Comparison of diverse developmental transcriptomes reveals that coexpression of gene neighbors is not evolutionarily conserved. Genome Res 19(12):2214-20.

Zaffran S, Astier M, Gratecos D, Semeriva M (1997) The held out wings (how) Drosophila gene encodes a putative RNA-binding protein involved in the control of muscular and cardiac activity. Development 124(10):2087-98.

Zhao D, McBride D, Nandi S, McQueen HA, McGrew MJ, Hocking P, Lewis PD, Sang H, Clinton M (2010) Somatic sex identity is cell autonomous in the chicken. Nature 464:237-242.

Zorn AM, Krieg PA (1997) The KH domain protein encoded by quaking functions as a dimer and is essential for notochord development in Xenopus embryos. Genes Dev 11(17):2176-90.