Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 349 (2005) 35–42
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Intraspecific variation in *fem-3* and *tra-2*, two rapidly coevolving nematode sex-determining genes

Eric S. Haag*, Amanda D. Ackerman

Department of Biology, University of Maryland, College Park, MD 20742, USA

Received 18 August 2004; received in revised form 16 December 2004; accepted 27 December 2004
Available online 17 March 2005
Received by D. Fitch

Abstract

The sex determination gene *fem-3* encodes one of the most divergent proteins yet described in the terrestrial nematode *Caenorhabditis*. Despite this rapid sequence change, however, FEM-3 is essential for male development in the three species surveyed thus far. It also participates in conserved protein–protein complexes with the transmembrane receptor TRA-2 and the phosphatase FEM-2 in these species. These interactions show strong species specificity, indicating that conserved residues are not sufficient for function and that compensatory evolution between binding partners is important. To shed further light on the nature of this coevolution, and to discern the extent of amino acid polymorphism allowed in FEM-3 and the domain of TRA-2 that binds it, we have examined intraspecific variation in the gonochoristic species *Caenorhabditis remanei*. Ten new complete *Cr-fem-3* alleles from three regions of the United States are described. We also obtained sequences for the FEM-3-binding domain of TRA-2 for 9 of the same strains. These alleles were compared with each other, with the European founder alleles, and with the orthologous sequences from the congeners *Caenorhabditis elegans* and *C. briggsae*. We find that FEM-3 harbors abundant amino acid polymorphisms along its entire length. The majority (but not all) of these occur in nonconserved residues, and in at least one domain there is evidence for diversifying selection. The FEM-3-binding domain of TRA-2 is less polymorphic than FEM-3. Amino acids neither polymorphic nor conserved between species are candidates for residues mediating species-specific interaction of FEM-3 with its binding partners.

Keywords: *C. remanei*; Polymorphism; Sex determination; Coevolution

1. Introduction

The regulation of sexual dimorphism is an important process in most animals, but the mechanisms by which it is accomplished are not conserved among phyla. For instance, sex chromosomes or environmental cues are two different ways in which animals may distinguish sexes. Even systems of genetic sex determination that are superficially similar may be completely different at the gene level (Cline and Meyer, 1996). Over smaller phylogenetic distances where conserved genes can be recognized, their sequences evolve rapidly compared to other genes in a diverse group of organisms (e.g., Tucker and Lundrigan, 1993; de Bono and Hodgkin, 1996; Civetta and Singh, 1998; Haag et al., 2002). Numerous studies in both *Drosophila* and in mammals have addressed whether this change is driven by adaptive evolution or by unconstrained neutral processes. Several of these studies have found that neutral evolution cannot be rejected despite the higher than usual substitution rates (O'Neil and Belote, 1992; McAllister and McVean, 2000; Moreira, 2002). Where positive selection has been inferred, it often fails to account for the bulk of sequence evolution (Pamilo and O'Neill, 1997; Jansa et al., 2003). Similar evaluation of evolutionary dynamics in nematode sex determination genes is hampered by the large genetic distance between presumably closely related species (Kiontke et al., 2004), including the mutational saturation of

Abbreviations: LD, linkage disequilibrium; PCR, polymerase chain reaction; EGF, epidermal growth factor; LNR, LIN-12/Notch repeat.

* Corresponding author. Tel.: +1 301 405 8534; fax: +1 301 314 9358.
E-mail address: ehaag@wam.umd.edu (E.S. Haag).

0378-1119/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2004.12.051

silent sites between congeners (Stein et al., 2003). As a result, many standard statistical methods to detect positive selection are not applicable.

In addition to selection pressures, the rapid evolution of sex determination genes also raises interesting questions about the relationship between primary protein sequences and biochemical function, and how molecular variation affects the developmental genetic network controlling sexual dimorphism. Here the nematode system fares better due to the extensive molecular characterization of the *C. elegans* sex determination gene network (Goodwin and Ellis, 2002) and recent interspecific comparisons of sex determination gene function (de Bono and Hodgkin, 1996; Haag and Kimble, 2000; Chen et al., 2001; Stothard et al., 2002; Haag et al., 2002; Pires-da Silva and Sommer, 2004). This study extends this work by examining within-species variation in the rapidly evolving nematode sex determination gene, *fem-3,* which is essential for all male somatic cell fates in *C. elegans*, *C. briggsae*, and *C. remanei* (Hodgkin, 1986; Haag et al., 2002). It encodes a novel cytoplasmic protein that acts genetically between *tra-2*, which encodes a membrane protein, and *tra-1*, which encodes a Gli-related zinc-finger transcription factor. FEM-3 interacts physically with both the cytoplasmic domain of TRA-2 (TRA-2c) and with another cytoplasmic protein, the PP2c-type phosphatase FEM-2 (Pilgrim et al., 1995; Chin-Sang and Spence, 1996; Mehra et al., 1999). In males, TRA-2A is inhibited by the secreted ligand HER-1, allowing FEM-3, along with FEM-1 and FEM-2, to repress *tra-1* activity and promote male development (reviewed by Kuwabara and Perry, 2001).

The physical interaction and epistatic relationship between *fem-3* and *tra-2* are conserved and necessary for proper male somatic development in the sister species *C. remanei* and *C. briggsae*. However, the FEM-3-binding domain (F3BD) of TRA-2c is hyperdivergent between species (Kuwabara, 1996; Mehra et al., 1999; Haag and Kimble, 2000), and only when FEM-3 and TRA-2c are from the same species can they interact successfully in yeast (Haag et al., 2002), implying the existence of compensatory coevolution. This phenomenon led us to examine the extent and location of allowed intraspecific polymorphisms in FEM-3 and the TRA-2c F3BD. If coevolution is characterized by bouts of strong selection, then little variation in either is expected to exist. Conversely, coevolution may be an essentially neutral process facilitated by a large amount of standing variation in the interacting proteins.

We also saw the potential to discern whether selection was driving divergence by comparing intraspecific variants from distinct locales. The gonochoristic species *C. remanei* was chosen as the organism of study because past research has shown that there is little genomic variation among populations of *C. elegans* (Graustein et al., 2002; Sivasundar and Hey, 2003). Sequencing of two *Ce-fem-3* alleles confirmed low sequence variation between populations (L. Phillips and E. Haag, unpublished data). The original *Cr-tra-2* and *Cr-fem-3* sequences (Haag and Kimble, 2000; Haag et al., 2002) were from SB146, the type strain isolated in Germany (Sudhaus and Kiontke, 1996). We have determined the entire sequence of 10 new alleles of *Cr-fem-3*, as well as 9 new alleles of the F3BD-encoding portion of *Cr-tra-2*, from populations in three regions of the United States. They reveal the existence of abundant polymorphic amino acids scattered along the length of *Cr*-FEM-3, and in one region we find evidence for their excess relative to the neutral expectation. The TRA-2c F3BD was considerably less variable. We further examine how the nonsynonymous polymorphisms relate to interspecies conservation and to previous functional studies of FEM-3. Silent changes are also abundant, including some large insertion/deletion polymorphisms.

2. Methods

C. remanei strains CR1014, CR1415, CR2124, EM464, PB205, PB206, PB212, PB219, PB228, and PB229 were obtained with the kind assistance of M. Palopoli, Bowdoin College, and cultured under standard *C. elegans* conditions (Wood, 1988) with the exception of increasing the agar content of plates to 2.5% to discourage burrowing. For each strain, genomic DNA was isolated as described in Haag and Kimble (2000) and used as template for the polymerase chain reaction (PCR) with primers annealing to the 5′ and 3′ untranslated portions of the terminal exons of *Cr-fem-3*, based on the sequence from strain SB146 (Haag et al., 2002). For strains refractory to this approach, *Cr-fem-3* was isolated as two overlapping fragments that together amplified the entire gene (Table 1). To minimize the chances of creating chimeric alleles, clones were assembled that were identical at polymorphic sites lying in the region of fragment overlap. When possible the entire *fem-3* coding region was amplified with Pfu and Taq polymerases.

Direct sequencing of the *Cr-fem-3* amplicons failed in many cases due to heterozygosity of indel polymorphisms, so most of the data were obtained from cloned material. Alignments of partial sequences from multiple clones from a given strain revealed multiple polymorphisms, similar to that reported in other studies (Haag and Kimble, 2000; Graustein et al., 2002). We therefore chose a single haplotype from each strain for complete characterization. Products larger than 900 base pairs were cloned into either the pCR-script (Stratagene) or the pCR 2.1 TOPO-TA (Invitrogen) plasmid vectors. Each gene was sequenced with a combination of direct sequencing of PCR products and sequencing of plasmid clones with a combination of GPS-1 transposon insertions (New England Biolabs), gene-specific primers, and flanking vector primers. Sequence reads were assembled into contigs with the GCG Wisconsin Package. Completed *C. remanei fem-3* alleles were aligned against