Available online at www.sciencedirect.com
SCIENCE DIRECT
Developmental Biology 289 (2006) 444 – 455
DEVELOPMENTAL BIOLOGY
www.elsevier.com/locate/ydbio

Genomes & Developmental Control

Med-type GATA factors and the evolution of mesendoderm specification in nematodes

Cristian Coroian, Gina Broitman-Maduro, Morris F. Maduro *

Department of Biology, UC Riverside, 3380 Spieth Hall, Riverside, CA 92521, USA

Received for publication 18 May 2005, revised 7 October 2005, accepted 22 October 2005
Available online 1 December 2005

Abstract

In the nematode, *C. elegans*, the divergent GATA-type transcription factors MED-1 and MED-2 are encoded by an unlinked, redundant pair of intronless genes. The *med-1,2* genes are among the first to be activated in the embryo and are critical for the specification of the 7-cell stage MS (mesoderm) and E (endoderm) precursor cells. We have previously shown that the binding site recognized by MED-1 is a noncanonical RAGTATAC site that is not expected from the resemblance of its single C4-type zinc finger to those of other known GATA factors, which recognize the consensus HGATAR. To date, no MED-like zinc fingers have been described outside of *C. elegans*. In order to understand the evolution of these transcription factors, and the evolution of gene networks that specify early cell fates in *Caenorhabditis*, we have identified *med* sequence homologs in the related nematodes *C. briggsae* and *C. remanei*. While *C. briggsae* encodes two *med*-like genes similar to *C. elegans*, we find evidence for seven distinct *med*-like genes in *C. remanei*. Somewhat unexpectedly, the coding regions of all *med* genes appear to lack introns. We report that the *med* homologs have similar expression in their respective species. We further show that the *C. briggsae* homologs, and at least five of the seven *C. remanei* homologs, can fully complement the embryonic lethal phenotype of a *C. elegans med-1,2(−)* strain. We conclude that Med function and expression have been conserved over tens of millions of years of evolution, and that there may be a mechanism that selects against the acquisition of introns in these genes.

Keywords: Intronless genes; *C. remanei*; med-1; Gene duplication; GATA factors; Zinc fingers; Transcription factors; Cell fate specification

Introduction

Deployment of appropriate gene regulatory networks is the means by which different cells acquire different fates during metazoan development (Levine and Davidson, 2005). Gene comparison among closely-related species is one method that can be used to learn more about such networks, as conserved features can reveal fundamental properties that change very little over evolutionary time (Yuh et al., 2002). Such studies contribute to our understanding of how the generation of diversity among metazoans is related to changes in the interactions among genes.

We are interested in the evolution of the gene regulatory network that operates in mesendoderm specification in *Caenorhabditis*. At the 7-cell stage of embryonic development in *C. elegans*, the precursor cells MS and E are born (Sulston et al., 1983). The E cell clonally establishes the entire intestine (midgut), which consists of 20 cells at hatching, while its sister cell, MS, generates 80 cells that are primarily mesodermal in character, including body wall muscle and cells of the posterior half of the feeding organ, the pharynx.

The transcriptional regulatory cascade that controls specification of MS and E in *C. elegans* is well understood (Fig. 1) (Maduro and Rothman, 2002). The maternal gene *skn-1* encodes a bZIP/homeodomain transcription factor required to specify MS and E: loss of *skn-1* function results in a transformation of both MS and E into the mesectodermal precursor C, which makes body wall muscle and hypodermis (Bowerman et al., 1992). In the mother of MS and E, called EMS, SKN-1 activates the expression of the *med-1* and *med-2* genes (Bowerman et al., 1993; Maduro et al., 2001). Depletion of *med-1,2* by RNAi also results in a transformation of MS and E to C, although E is correctly specified in approximately half of mutant embryos (Maduro et al., 2001). In the E cell, MED-1,2 activate the genes *end-1* and *end-3*, which encode a pair of

* Corresponding author. Fax: +1 951 827 4286.
E-mail address: mmaduro@citrus.ucr.edu (M.F. Maduro).

0012-1606/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2005.10.024

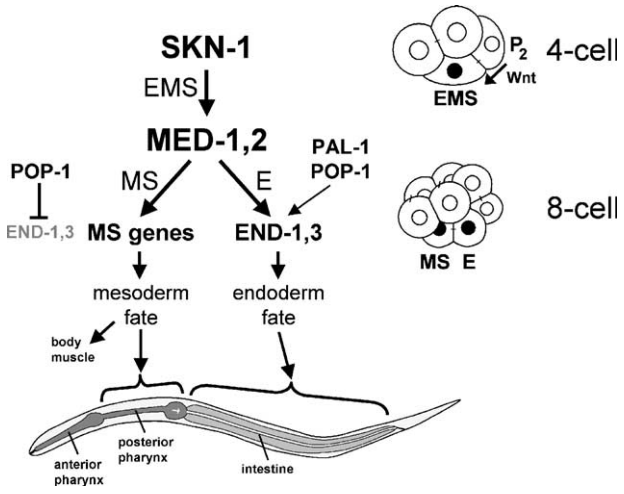

Fig. 1. Simplified representation of the regulatory cascade specifying MS and E in *C. elegans*. Diagrams of a 4-cell and 8-cell embryo are shown for reference. The maternal SKN-1 transcription factor activates zygotic expression of *med-1* and *med-2* in EMS (Maduro et al., 2001). In turn, MED-1,2 activate target genes in MS and E to specify their fates. In the E cell, MED-1,2 activate *end-1,3* to specify an endodermal fate (Maduro et al., 2002). The Wnt effector POP-1 is a coregulator along with the MEDs: in MS, POP-1 represses *end-1,3*, while in the E cell, the repressive activity of POP-1 is blocked as a result of the Wnt signaling event that polarizes EMS at the 4-cell stage (Goldstein, 1992; Maduro et al., 2002; Rocheleau et al., 1997; Thorpe et al., 1997). In E, POP-1 and PAL-1 also function as *end-1,3* activators in parallel with the MEDs (Maduro and Rothman, 2002; Maduro et al., 2005b). MS gives rise to primarily mesodermal cell types, including body wall muscle and the posterior half of the pharynx. The E cell clonally establishes the entire midgut (intestine). Sister cells are joined by a short line, and the nuclei of EMS, MS, and E are shaded black. In this and other figures, embryos are shown with anterior to the left and dorsal up.

GATA factors that specify an endoderm fate (Maduro and Rothman, 2002; Maduro et al., 2005a). In the MS cell, the corepressor POP-1 blocks the ability of MED-1,2 to activate *end-1,3*, allowing the *med*s to promote a mesodermal fate (Maduro et al., 2002). The repressive function of POP-1 is blocked in the E cell by transduction of a Wnt/MAPK/Src signal that originates from a cell–cell interaction between EMS and its sister cell, P_2 (Bei et al., 2002; Goldstein, 1992; Rocheleau et al., 1997, 1999; Shin et al., 1999; Thorpe et al., 1997). POP-1 then functions as an activator, and along with the Caudal-like protein PAL-1, also contributes to E specification in parallel with *med-1,2* (Maduro et al., 2005b). The *med-1* and *med-2* genes therefore play a critical role in the specification of the MS and E progenitors, as they function at the interface of SKN-1 activation and transduction of a signal initiated by the P_2–EMS interaction.

The *med-1* and *med-2* genes encode nearly-identical single C4 zinc finger proteins with similarity to the GATA family of transcription factors, which recognize the canonical binding site HGATAR (Lowry and Atchley, 2000; Maduro et al., 2001). MED-1 (and presumably MED-2, whose DNA-binding domain differs by only a single amino acid) recognizes RAGTATAC sites, suggesting that MED-1,2 have diverged in function from the GATA family (Broitman-Maduro et al., 2005). No proteins similar to MED-1 are known to exist outside of *Caenorhabditis* that are not themselves more similar to canonical GATA factors, suggesting that they may be a nematode invention. Because of the uniqueness of the MEDs among eukaryotic regulators, little is known of their evolutionary origin, and structure–function properties of the MED DNA-binding domain have yet to be elucidated.

In order to expand the collection of known MED-like regulators, and as first step toward studying the mesendoderm gene network in other species, we have mined the genome sequences of the related nematodes, *C. briggsae* (a hermaphroditic species) and *C. remanei* (a male–female species). *C. remanei* and *C. briggsae* are more closely related to each other than *C. elegans* is from either (Cho et al., 2004; Kiontke et al., 2004). By some estimates, *C. elegans* is 80–110 million years diverged from the last common ancestor to *C. briggsae*, and hence *C. remanei* (Coghlan and Wolfe, 2002; Stein et al., 2003). In this study, we compare the structure and genome organization of the *med* homologs from *C. briggsae* (two *med* genes) and *C. remanei* (seven *med* genes), and provide evidence that their expression and function have been conserved.

Materials and methods

Strains and genetics

The following strains were used in this work: PB4641 (wild-type *C. remanei* strain), AF16 (wild-type *C. briggsae*), N2 (wild-type *C. elegans*), MS240 [*unc-119(ed4) III; irEx102[unc-119(+), pMM808]*], MS247 [*med-1(ok804) X; med-2(cx9744) III; irDp1(III;f)*]. Transgenic strains used to test rescue are listed in Table 1. The free duplication *irDp1* (carried by MS247) was derived from the free duplication *sDp3*, which carries a single copy of *med-2*, after spontaneous integration of an array (*irEx14*) consisting of the *Ce-med-1* plasmid pMM277, the *unc-32*-rescuing plasmid pAIE5 (Pujol et al., 2001), and an *unc-119*::YFP reporter plasmid (pMM531). A more detailed description of the properties of MS247 will appear elsewhere.

Table 1
Rescue of a *C. elegans med-1,2(−)* mutant with *med* homolog transgenes

Strain(s)	Transgene	Rescue[a]	Viable	Arrested embryo		Total
				With gut	Without gut	
MS247	None[b]	n/a	0% (0)	58% (70)	42% (50)	120
MS290, MS291	*Ce-med-1*	Yes	97% (203)	1% (3)	1% (3)	209
MS402, MS403	*Ce-med-2*	Yes	98% (353)	2% (6)	0% (1)	360
MS298, MS299	*Cb-med-1*	Yes	80% (237)	18% (53)	3% (8)	298
MS420	*Cb-med-2*	Yes	33% (14)	57% (24)	10% (4)	42
MS293, MS295	*Cr-med-1*	Yes	76% (148)	24% (46)	0% (0)	194
MS335, MS336	*Cr-med-2*	No	0% (0)	61% (57)	39% (37)	94
MS314, MS315	*Cr-med-3*	Yes	86% (128)	13% (19)	1% (2)	149
MS317, MS393	*Cr-med-4*	Yes	64% (319)	35% (176)	1% (7)	502
MS319, MS320	*Cr-med-5*	Yes	87% (290)	11% (38)	2% (5)	333
MS322, MS323	*Cr-med-6*	Yes	93% (169)	6% (11)	1% (2)	182
MS337, MS338	*Cr-med-7*	No	0% (0)	90% (53)	10% (6)	59

[a] MS247 [*med-1(ok804); med-2(cx9744); irDp1*] animals were made transgenic for a plasmid or PCR product containing the indicated gene and an *unc-119*::CFP reporter. Lines that could be propagated in the absence of the *irDp1* balancer were considered to be rescued. For those that did not rescue, lines were constructed that also contained the *rol-6*^{D} marker plasmid (pRF4) to facilitate maintenance of the transgenes. Only animals expressing *unc-119*::CFP (indicative of the test array) but not *unc-119*::YFP (indicative of *irDp1*) were scored. For those genes that rescued, only *unc-119*::CFP embryos were scored.

[b] Non-*irDp1* MS247 segregants (i.e., embryos lacking *unc-119*::YFP) were scored. n/a, not applicable.