Developmental Biology ■ (■■■■) ■■■–■■■

Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/locate/developmentalbiology

Mutagenesis of GATA motifs controlling the endoderm regulator *elt-2* reveals distinct dominant and secondary *cis*-regulatory elements

Lawrence Du, Sharon Tracy, Scott A. Rifkin *

Section of Ecology, Behavior, and Evolution, Division of Biological Sciences, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0116, United States

ARTICLE INFO

Article history:
Received 1 December 2015
Received in revised form
3 February 2016
Accepted 10 February 2016

Keywords:
Endoderm
Cis-regulation
Promoter
smFISH
Transcription
GATA
C. elegans

ABSTRACT

Cis-regulatory elements (CREs) are crucial links in developmental gene regulatory networks, but in many cases, it can be difficult to discern whether similar CREs are functionally equivalent. We found that despite similar conservation and binding capability to upstream activators, different GATA *cis*-regulatory motifs within the promoter of the *C. elegans* endoderm regulator *elt-2* play distinctive roles in activating and modulating gene expression throughout development. We fused wild-type and mutant versions of the *elt-2* promoter to a *gfp* reporter and inserted these constructs as single copies into the *C. elegans* genome. We then counted early embryonic *gfp* transcripts using single-molecule RNA FISH (smFISH) and quantified gut GFP fluorescence. We determined that a single primary dominant GATA motif located 527 bp upstream of the *elt-2* start codon was necessary for both embryonic activation and later maintenance of transcription, while nearby secondary GATA motifs played largely subtle roles in modulating postembryonic levels of *elt-2*. Mutation of the primary activating site increased low-level spatiotemporally ectopic stochastic transcription, indicating that this site acts repressively in non-endoderm cells. Our results reveal that CREs with similar GATA factor binding affinities in close proximity can play very divergent context-dependent roles in regulating the expression of a developmentally critical gene *in vivo*.

1. Introduction

For the nematode *C. elegans* and its close relatives, early embryonic development is characterized by a tight link between cell lineage and cell fate that is largely determined by transcriptional gene regulatory networks (GRNs). Determining how transcription factors activate their respective targets within a GRN at the *cis*-regulatory level is key to understanding how multicellular organisms develop robustly.

However, understanding *cis*-regulation has been complicated by the fact that the vast majority of eukaryotic transcription factors have very short DNA binding domains, often leading to vastly more potential cognate binding sites than real functional targets (Mirny and Wunderlich, 2008). Additionally, individual transcription factors can co-occur with one or more paralogous factors from the same family leading to multiple transcription factors sharing individual *cis*-regulatory sites. Although recent years have seen vast advances in the mapping of transcription factors to their binding sites though techniques such as ChIP-seq (Gerstein et al., 2010), such techniques do not necessarily reveal whether bound sites are functionally equivalent. Many transcription factors are known to act as both activators and repressors depending on context.

Low target specificity, gene duplications, and contextual role switching have the potential to play a role in advancing developmental robustness. During early embryonic development, transcriptional networks must be robust to extrinsic insults as well as intrinsic variability at the molecular level. Cell divisions need to be spatially and temporally coordinated in the face of environmental variability and stochastic fluctuations of key molecules.

The transcriptional regulation of the *C. elegans* endoderm specifying gene *elt-2* is a good model for studying how *cis*-regulatory mechanisms impact developmental robustness. The gene *elt-2* is an essential switch for the endoderm cell fate decision and a fundamental developmental bottleneck: failure to activate *elt-2* results in a lethal absence of endoderm. The major *trans*-activators of *elt-2* are well characterized and have been demonstrated to contribute to developmental robustness at the *trans*-level. END-3, END-1, and ELT-7 are closely related GATA transcription factors that redundantly activate *elt-2* during early embryonic development (Lowry et al., 2009; Raj et al., 2010; Sommermann et al., 2010; Zhu et al., 1997, 1998), and ELT-2 maintains its own transcription through larval and adult stages by autoregulation

* Corresponding author.
E-mail address: sarifkin@ucsd.edu (S.A. Rifkin).

http://dx.doi.org/10.1016/j.ydbio.2016.02.013
0012-1606/

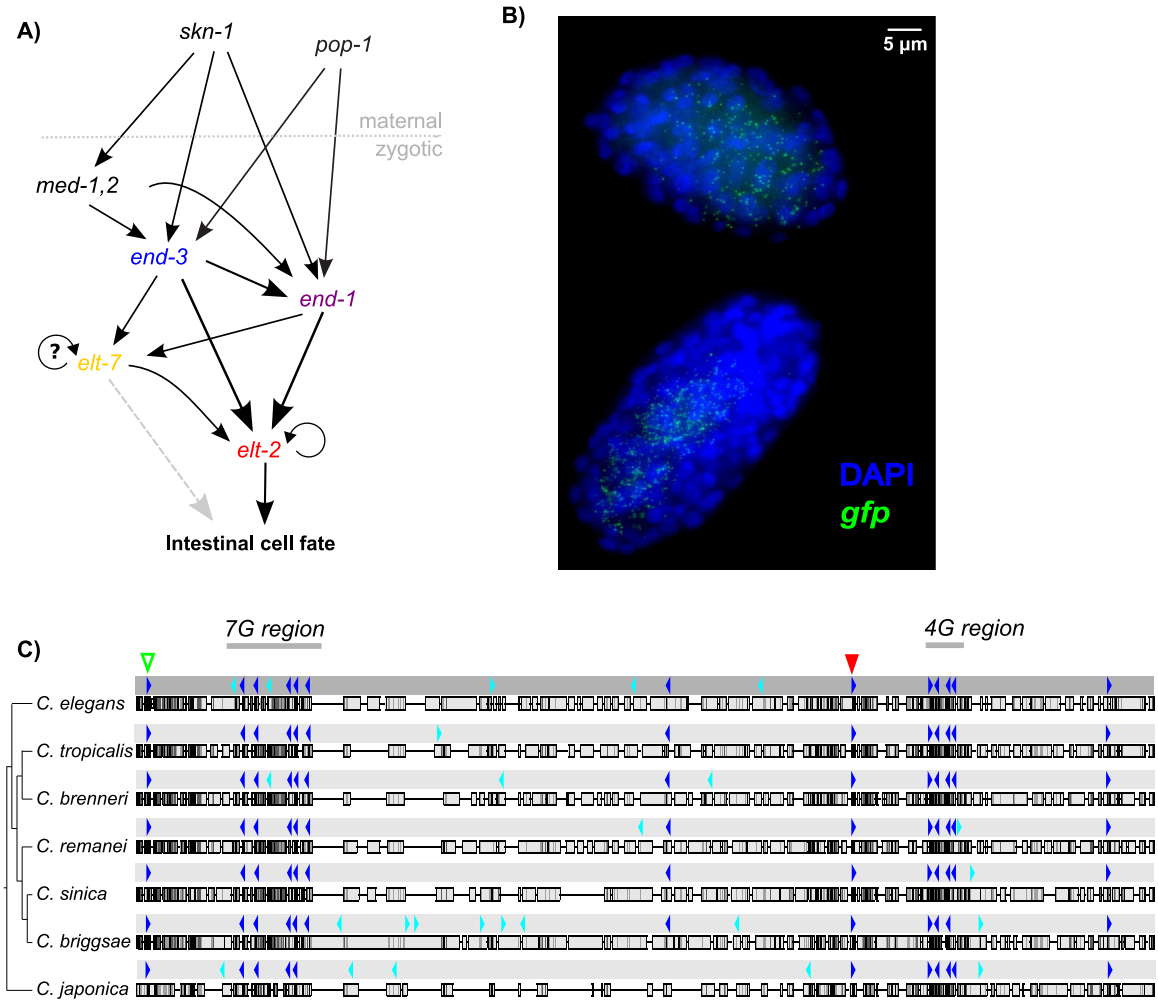

Fig. 1. The *elt-2* activation network and upstream promoter region. (A) Network diagram of *elt-2* activators. (B) Maximum Z-stack projection of an smFISH image from strain 1879:WT. (C) 1879 bp promoter region upstream of *elt-2* aligned to orthologous regions in *Elegans* supergroup members. Dark blue indicates a heavily conserved HGATAR motif and cyan indicates a weakly conserved HGATAR motif. The red, filled triangle indicates the position of the −527 bp ACTGATAAGA "A-site" motif. The green, open triangle indicates position of a −1857 bp ACTGATAAGG motif. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(Fukushige et al., 1999) (Fig. 1A). Single null mutants of END-1, END-3, or ELT-7 are largely viable-with only transient developmental anomalies occurring in *end-1* and *end-3* single null mutants and a low (5–9%) rate of developmental failure in *end-3* mutants (Boeck et al., 2011; Maduro et al., 2005; Sommermann et al., 2010). A paralogous pair of redundant and nearly identical GATA factors-*med-1,2*-also helps to activate *end-3* and *end-1* (Maduro et al., 2001, 2007). Raj et al. (2010) demonstrated that particular mutations in the upstream maternal activating factor *skn-1* result in failure to activate *med-1,2* and *end-3* and highly variable expression of *end-1*. Noisy *end-1* expression, in turn, leads to bimodal *elt-2* expression states (Raj et al., 2010). The presence of redundant *trans*-activating factors effectively buffers the activation of *elt-2* from variability in levels of any single activator.

Despite our good understanding of *elt-2*'s *trans*-activators, little is known about how these *trans*-activators operate at the *cis*-regulatory level. The exact sequences and relative positions of the *cis*-regulatory motifs necessary for driving *elt-2* expression have not been determined, nor are there apparent TATA box (GTATAWWAG) or Sp1 core promoter motifs in the immediate region upstream of the *elt-2* transcriptional start site (TSS) (WormBase release WS220) (Harris et al., 2010; Saito et al., 2013). Furthermore, we could not identify any sequence similarities to the known basal promoter fragment of the *pes-10* gene. Based on the fact that all the known embryonic activators of *elt-2* are GATA transcription factors, we can narrow down the candidate *cis*-regulatory sequences considerably. The ~2 kb region upstream of the *elt-2* start codon in *C. elegans* contains 18 of the consensus GATA factor binding motif HGATAR, with 13 conserved in sequence and relative spacing throughout the *Elegans* supergroup (sequence data from *C. elegans, C. tropicalis* , *C. brenneri*, *C. remanei*, *C. sinica*, *C. briggsae*, and *C. japonica*) (Félix et al., 2014; Huang et al., 2014) (Fig. 1C).

This conglomeration of potential GATA factor binding sites in the *elt-2* upstream region suggests several possible ways that these motifs might interact with the *trans*-factors END-3, END-1, ELT-7, and ELT-2 itself to control *elt-2* transcriptional activation. Perhaps many independently dispersed transcription start sites, driven or aided by GATA factor binding, contribute additively and redundantly to overall gene expression levels and noise (Juven-Gershon and Kadonaga, 2010). Under this dispersed promoter scenario, mutation of single HGATAR motifs might be expected to reduce transcription activation proportionate to the number of motifs mutated (Davidson, 2001; Flores et al., 2000). Alternatively, but not exclusively, *elt-2 cis*-activation could also be driven by a combinatorial code involving binding of different GATA factors with different specificities. Under a combinatorial control scenario, mutation of any single HGATAR motif in a larger combinatorial code should result in an equivalent impact on gene expression as mutating any single motif in the same code. Finally, the contributions of different sites may instead be unequal, with one or a

Please cite this article as: Du, L., et al., Mutagenesis of GATA motifs controlling the endoderm regulator *elt-2* reveals distinct dominant and secondary *cis*-regulatory elements. Dev. Biol. (2016), http://dx.doi.org/10.1016/j.ydbio.2016.02.013