A developmental gene regulatory network for invasive differentiation of the *C. elegans* anchor cell

Taylor N. Medwig-Kinney*, Jayson J. Smith*, Nicholas J. Palmisano, Wan Zhang and David Q. Matus‡

Department of Biochemistry and Cell Biology, Stony Brook University, Stony Brook, NY 11794-5215, USA.

*These authors contributed equally to this work

‡Author for correspondence (david.matus@stonybrook.edu)

SUMMARY STATEMENT

Basement membrane invasion by the *C. elegans* anchor cell is coordinated by a gene regulatory network encompassing overlapping cell cycle dependent and independent sub-circuits.

ABSTRACT

Cellular invasion through the basement membrane is a key part of fertility, development, immunity, and disease. To understand the acquisition of the invasive phenotype, we use the genetically tractable *in vivo* model of *C. elegans* anchor cell invasion into the vulval epithelium. Coordination of AC invasive machinery is largely achieved cell autonomously through the activity of four conserved transcription factors (TFs), *fos-1a* (Fos), *egl-43* (EVI/MEL), *hlh-2* (E/Daughterless) and *nhr-67* (NR2E1/TLX). Here, using genome editing and protein depletion with improved RNA interference (RNAi), we characterize the gene regulatory network (GRN) programming AC invasion. Using quantitative measures of endogenous protein levels paired with perturbation analyses, we identify the upstream regulation of *nhr-67* activity by *egl-43* and *hlh-2*. Network inference suggests that these three TFs function in a type I coherent feed-forward loop with positive feedback to maintain the AC in a post-mitotic, pro-invasive state. Finally, we provide evidence that two overlapping pro-invasive sub-circuits function in a cell cycle-dependent and independent fashion. Together, these results provide the first framework for integrating TF relationships with the cell biological mechanisms driving invasive cellular behavior.

Key words: EGL-43, FOS-1, HLH-2, NHR-67, gene regulatory network, cell invasion

INTRODUCTION

Invasion through basement membranes (BM) is a cellular behavior integral to the establishment of pregnancy, tissue patterning during embryonic development, and immune response to infection and injury (Medwig and Matus, 2017; Rowe and Weiss, 2008). Increased cellular invasiveness is also a hallmark of metastatic cancer (Hanahan and Weinberg, 2011). Previous research has identified several cell-autonomous mechanisms that are highly conserved across different contexts of BM invasion (Kelley et al., 2014; Medwig and Matus, 2017). These include expression of extracellular matrix (ECM) proteins to facilitate cell-BM adhesion, localization of F-actin into invasive protrusions that physically puncture the BM, as well as upregulation of matrix metalloproteinases (MMPs) to chemically degrade the BM (Hagedorn et al., 2009; Hagedorn et al., 2013; Kelley et al., 2018; Morrissey et al., 2014; Sherwood et al., 2005). There is also growing evidence that cells must undergo cell cycle arrest in order to achieve invasive differentiation (Kohrman and Matus, 2017; Matus et al., 2015). How these tightly coordinated programs are transcriptionally regulated is not well understood.

As many contexts of cellular invasion occur deep within tissue layers where it is difficult to visualize, we utilize morphogenesis of the *C. elegans* uterine-vulval connection as a genetically tractable and visually amenable model for examining cell invasion *in vivo*. During the mid-L3 stage, a specialized uterine cell called the anchor cell (AC), invades through the underlying BM in order to connect the uterus to the vulval epithelium, facilitating egg-laying (Sherwood and Sternberg, 2003). The AC itself is specified in a cell fate decision event earlier in development, during the L2 stage, where two equipotent cells diverge via stochastic Notch asymmetry, giving rise to the terminally differentiated AC and a proliferative ventral uterine (VU) cell (Wilkinson et al., 1994).

Prior research has identified four pro-invasive transcription factors (TFs) that function cell autonomously to regulate AC invasion (Fig. 1). These include the basic leucine zipper TF *fos-1a* (Fos), the basic helix-loop-helix TF *hlh-2* (E/Daughterless), the nuclear hormone receptor *nhr-67* (NR2E1/Tailless/TLX) and the zinc-finger TF *egl-43*

(EVI1/MEL1) (Hwang et al., 2007; Matus et al., 2010; Rimann and Hajnal, 2007; Schindler and Sherwood, 2011; Sherwood et al., 2005; Verghese et al., 2011). NHR-67 functions upstream of the cyclin dependent kinase inhibitor, CKI-1 (p21/p27), to induce G1/G0 cell cycle arrest, which is necessary for AC invasion (Matus et al., 2015). Independent of NHR-67 activity, FOS-1A regulates the expression of the long isoform of *egl-43* (EGL-43L) and downstream effectors including multiple MMPs (*zmp-1, -3, and -6*), a cadherin (*cdh-3*), and hemicentin (*him-4*), an immunoglobulin superfamily protein that generates a linkage between the ventral epidermal and somatic gonad BMs during invasion (Hwang et al., 2007; Kelley et al., 2018; Matus et al., 2010; Morrissey et al., 2014; Rimann and Hajnal, 2007). HLH-2 independently regulates *cdh-3* and *him-4* as well as an ECM glycoprotein (MIG-6/papilin) and cytoskeletal polarity (Schindler and Sherwood, 2011). Prior work has additionally suggested that EGL-43 and HLH-2 may regulate NHR-67 based on binding motifs present in the *nhr-67* promoter (Bodofsky et al., 2018; Verghese et al., 2011). How these four TFs function, both independently and together, to regulate the invasive activity of the AC is poorly understood.

Here, using new highly efficient RNA interference (RNAi) vectors (Sturm et al., 2018), we identify previously unreported AC invasion phenotypes associated with depletion of *egl-43* and *hlh-2*. In addition, using CRISPR/Cas9-genome engineering, we have endogenously GFP-tagged the genomic locus of each TF, allowing us to report both the relative expression of pro-invasive TFs in the AC and dissect their molecular epistatic interactions. Based on these findings, we have characterized cell cycle dependent and independent pathways necessary for invasion. Finally, we have identified a feed-forward regulatory circuit with positive feedback critical for maintaining the AC in a pro-invasive, post-mitotic state. These findings provide new insights into the GRN that underlies invasive cell behavior.

RESULTS

Improved RNAi penetrance reveals new phenotypes associated with depletion of pro-invasive TFs

C. elegans AC invasion is a simple, tractable *in vivo* model for understanding how cells breach the barriers imposed by BM (Gupta et al., 2012; Sherwood and Plastino, 2018). We and others have identified four TFs that function cell autonomously in the AC to regulate invasion (Hwang et al., 2007; Matus et al., 2015; Rimann and Hajnal, 2007; Schindler and Sherwood, 2011; Sherwood et al., 2005). Recently, we identified that the post-mitotic AC must be in a G1/G0-cell cycle arrested state, regulated by *nhr-67,* in order to adopt an invasive phenotype (Fig. 1) (Matus et al., 2015). Roles for the other three TFs in AC specification (*egl-43* and *hlh-2*) (Hwang et al., 2007; Karp and Greenwald, 2004; Rimann and Hajnal, 2007; Sallee et al., 2017; Verghese et al., 2011) and the transcriptional regulation of downstream invasive effectors, including MMPs *(fos-1a* and *egl-43*) and other pro-invasive targets, have been previously established (Fig. 1A) (Bodofsky et al., 2018; Kelley et al., 2018; Matus et al., 2010; Matus et al., 2015; Schindler and Sherwood, 2011; Sherwood et al., 2005; Wang et al., 2014). Here, we sought to gain a deeper understanding of the regulatory relationships between pro-invasive TFs in the AC, taking advantage of recent advances in RNAi targeting (Sturm et al., 2018) and CRISPR/Cas9-genome engineering (Dickinson and Goldstein, 2016). To accomplish this, we first generated new RNAi targeting constructs in the improved RNAi vector, T444T, which includes T7 termination sequences to prevent the generation of non-specific RNA fragments from the vector backbone, increasing the efficacy of gene silencing over the original RNAi targeting vector, L4440 (Sturm et al., 2018).

As three of four pro-invasive TFs also function during AC specification (Hwang et al., 2007; Karp and Greenwald, 2004; Verghese et al., 2011), we scored TF depletions in a genetic background where only the AC and neighboring uterine cells are sensitive to RNAi. The uterine-specific RNAi sensitive strain was generated through tissue-specific restoration of the RDE-1 Piwi/Argonaut protein in an *rde-1* mutant background using the *fos-1a* promoter, which is expressed specifically in the somatic gonad during the late L2/early L3 stage of larval development (Haerty et al., 2008; Hagedorn et al., 2009; Matus et al., 2010; Matus et al., 2015). AC invasion was quantified as the presence or absence of a BM gap, visualized by laminin::GFP, at the P6.p four-cell stage, a

developmental window where 100% of wild-type ACs are invaded (Fig. 2). We also scored invasion several hours later, following vulval morphogenesis, allowing us to distinguish between delays in invasion and complete loss of invasive capacity (Fig. 2B, Fig. S1, Table S1). Depletion of each TF resulted in invasion defects ranging from moderate penetrance (38%, *fos-1*) to highly penetrant (69-94%, *hlh-2*, *egl-43*, and *nhr-67*), when synchronized L1-stage animals were plated on RNAi-containing bacteria and scored at the mid-L3 stage (P6.p 4-cell stage) and early L4 stage (P6.p 8-cell stage) (Fig. 2, Fig. S1). Consistent with previous work, depletion of *fos-1* resulted in single ACs that failed to breach the BM. Depletion of *nhr-67* resulted in proliferative, non-invasive ACs at high penetrance. Interestingly, depletion of *egl-43* phenocopied *nhr-67*, with a highly penetrant defect of multiple, non-invasive ACs. Lastly, loss of *hlh-2* resulted in a range of phenotypes, from animals with zero, one, two or even several cases of three or more ACs (Fig. S1B,C). In all cases, the penetrance of AC invasion defects was significantly increased using a T444T-based RNAi targeting vector as compared to L4440-based vectors (Fig. S1A). Thus, use of the improved RNAi vector revealed potentially novel phenotypes for depletion of *egl-43* and *hlh-2* during invasion.

As we did see a range of phenotypes in *hlh-2* depleted animals, including animals possessing multiple ACs, we wanted to rule out any potential defects in AC specification. Depletions using the traditional RNAi vector, L4440, in the uterine-specific RNAi-sensitive background is sufficient to bypass AC fate specification defects that occur in genetic backgrounds where all cells are sensitive to RNAi (Matus et al., 2010). Specifically, loss of HLH-2 prior to AC specification regulates the *lin-12*(Notch)/*lag-2*(Delta)-dependent signaling cascade, resulting in both pre-AC/VU cells adopting a VU fate. Later, loss of HLH-2 during AC specification results in upregulation of pro-AC fate by activating *lag-2*/Delta resulting in both AC/VU cells adopting an AC fate (Karp and Greenwald, 2004; Schindler and Sherwood, 2011). Notably, *hlh-2* depletion using the more penetrant T444T vector resulted in a significantly greater occurrence of a 0 AC phenotype as compared to an L4440 *hlh-2* RNAi clone (Fig. S1A-C). To bypass specification defects due to increased RNAi penetrance for *hlh-2* using the T444T vector, we depleted *hlh-2* by RNAi using an L2-plating strategy, growing animals to the

time of the L1/L2 molt and then transferring them to *hlh-2(RNAi)* plates. The L2 plating resulted in a lower penetrance overall of AC invasion defects (43% at the P6.p 4-cell stage; Fig. 2, Fig. S1C), but all animals possessed an AC. Strikingly, multiple animals possessed multiple *cdh-3*-expressing ACs early (Fig. 2), while 10% of animals in the early L4 stage possessed multiple ACs (Fig. 2B, Fig. S1C,D). As the number of ACs increased over developmental time (Fig. 2B, Fig. S1C,D), these results suggest that the multiple ACs could be arising through loss of G1/G0-cell cycle arrest and *nhr-67* activity (Matus et al., 2015).

Generation of GFP-tagged proteins to establish a quantitative framework for examining TF interactions

Next, we sought to examine the relationship of these four TFs in the AC prior to and during invasion. To do this in a quantitative fashion, we used CRISPR/Cas9-genome editing technology to knock-in a codon-optimized GFP tag into the endogenous locus of each TF (Fig. S2) (Dickinson and Goldstein, 2016; Dickinson et al., 2013; Dickinson et al., 2015). All four TFs showed robust GFP-localization in the AC nucleus prior to and during invasion (Fig. 3A). Additionally, the GFP-tagged strains had similar expression domains in the somatic gonad during the L3 larval stage as previously reported by traditional multi-copy array transgenes (Hwang et al., 2007; Matus et al., 2015; Rimann and Hajnal, 2007; Sherwood et al., 2005; Verghese et al., 2011). After examining expression patterns, we then synchronized TF GFP-tagged strains and collected a developmental time-course quantifying the expression of GFP-tagged protein in the AC at the P6.p 1-cell, 2-cell, 4-cell and 8-cell stages (Fig 3A). All GFP-tagged TFs were imaged using uniform acquisition settings on a spinning disk confocal using an EM-CCD camera, allowing us to compare relative expression levels of endogenous GFP-tagged proteins. Interestingly, all four TFs followed the same general trend in expression levels, with a gradual increase in levels prior to invasion, peaking at or just before the P6.p 4-cell stage when 100% of wild-type animals have generated a gap in the BM. Expression levels then decreased following invasion, during vulval morphogenesis in the early L4 stage (P6.p 8-cell stage) (Fig. 3B).

Correlating phenotypes to quantitative loss of endogenous TFs

We next wished to assess the efficacy of our newly generated TF-RNAi vectors quantitatively. To accomplish this, we first crossed a BM reporter (laminin::GFP) and an AC marker (*cdh-3>mCherry:moeABD*) into each endogenously GFP-tagged TF strain. We then performed a series of RNAi depletion experiments targeting each TF and examining synchronized animals at the P6.p 4-cell stage. As the insertion of GFP-tags into native genomic loci can potentially interfere with gene function, we examined a minimum of 50 animals treated with control (T444T) empty vector RNAi (Fig. 4). All four GFP-tagged strains showed 100% BM breach at the normal time of invasion in control animals (empty vector) (Fig. 4). Next, we examined a minimum of 50 animals following targeted TF-RNAi depletions, collecting spinning disk confocal z-stacks using the same acquisition settings as used for the developmental series (Fig. 3) in all experiments. Following image quantification, our results were consistent with the phenotypes identified in our original uterine-specific RNAi screen (Fig. 2,4). However, by being able to quantify loss of endogenous GFP-tagged TF targeted by RNAi, we were more accurately able to correlate phenotype with protein depletion. The *nhr-67* and *egl-43* improved RNAi constructs strongly knocked down their GFP-tagged endogenous targets, with depletions averaging 88% and 75%, respectively (Fig. 4E). This strong knockdown of *nhr-67* and *egl-43* was also correlated with a highly penetrant AC invasion defects (61% and 100%, respectively). The completely penetrant defect from *egl-43(RNAi)* segregated into 26% of animals exhibiting a single AC that failed to invade and the remaining phenocopying loss of *nhr-67*, with multiple *cdh-3*-expressing non-invasive ACs (Fig. 4A,E). RNAi depletion of *GFP::fos-1a* was also penetrant, with a mean depletion of 92% and 76% of treated animals exhibiting a block in AC invasion (Fig. 4B,E). To assess post-AC specification phenotypes following *hlh-2* knockdown, we performed L2 platings and scored for defects. As the animals experience ~12 hours (at 25°C) less time exposed to *hlh-2(RNAi)*, we were not surprised to see relatively weaker *GFP::hlh-2* depletion as compared to L1 platings targeting the other TFs (Fig. 4E). However, this shorter time of depletion revealed a strong correlation between invasion

defect and percent protein depletion. Animals with ~50% depletion appeared wild-type, with a single AC that breached the BM. Animals with stronger depletion in the AC presented pleiotropic phenotypes, much like *egl-43(RNAi)* (Fig. 4C,E). Notably, the average depletion for animals with a single non-invasive AC was weaker than for animals with multiple *cdh-3*-expressing non-invasive ACs (60% vs. 78% respectively, Fig. 4D,E). These results suggest that following strong TF depletion, loss of either *egl-43* or *hlh-2* phenocopy *nhr-67* depletion, generating multiple, non-invasive ACs.

Identification of a feed-forward loop controlling NHR-67 activity

Next, we examined the relationship between the four pro-invasive TFs during invasion. To accomplish this, we performed a series of RNAi depletion experiments using synchronized L1 or L2 (for *hlh-2*) stage animals and quantified the amount of GFP-tagged endogenous TF in the AC in animals with defects in invasion. While depletion of *fos-1* failed to significantly down-regulate levels of *nhr-67::GFP,* loss of *egl-43* resulted in a strong reduction (65%) of *nhr-67::GFP* in the AC (Fig. 5A,B). Intriguingly, in animals with a single non-invasive AC following *hlh-2(RNAi),* we saw only a partial reduction (19%) of *nhr-67::GFP* levels, while in animals with multiple *cdh-3*-expressing cells we saw a significant reduction (49%) in *nhr-67::GFP* (Fig. 5A,B). These results suggest that both *egl-43* and *hlh-2* co-regulate *nhr-67* during invasion.

We repeated these TF-RNAi molecular epistasis experiments with the remaining three TF-GFP-tagged strains and quantified loss of GFP in animals with invasion defects. Consistent with previous studies using transcriptional GFP promoter reporters (Hwang et al., 2007; Rimann and Hajnal, 2007), we found that depletion of *fos-1* regulated levels of *egl-43::GFP::egl-43* (44% depletion) (Fig. 5C,D). No other TF depletion significantly regulated the levels of *egl-43* in our experiments (Fig. 5C,D). *hlh-2* is predicted to regulate *egl-43* activity based on the presence of two conserved E-box binding motifs in the *egl-43* promoter (Hwang et al. 2007). We did not detect regulation of *egl-43::GFP::egl-43* in either the AC (Fig. 5C,D) or neighboring ventral uterine (VU) cells (Fig. S3), but as our experiments with *hlh-2* depletion were done as L2 platings post-AC

specification, it is possible that *hlh-2* might regulate *egl-43* prior to AC/VU specification. Next, we examined *GFP::hlh-2* levels following TF-RNAi depletions. Intriguingly, we found a similar pattern of regulation based on AC phenotype. Animals with single non-invasive ACs showed partial depletion of *GFP::hlh-2* following *fos-1, egl-43,* or *nhr-67* loss (35%, 20%, and 14%, respectively). However, animals with multiple *cdh-3*-expressing ACs following *egl-43(RNAi)* or *nhr-67(RNAi)* showed strong depletion of *GFP::hlh-2* (66% and 73%, respectively) (Fig. 5E,F). Finally, we examined levels of *GFP::fos-1a* following TF depletions. Loss of either *nhr-67* or *egl-43* resulted in partial reduction of *GFP::fos-1a* levels in the AC (52% and 49%, respectively) (Fig. 5G,H). Together, these results suggest that *egl-43*, *hlh-2* and *nhr-67* may function together in a feed-forward regulatory loop with positive feedback to maintain the AC in a post-mitotic, pro-invasive state.

Multiple ACs derive from proliferation

As loss of *egl-43* and *hlh-2* both regulate levels of *nhr-67* and their depletion results in multiple ACs that fail to invade, we next wanted to assess whether their depletion was functionally phenocopying loss of *nhr-67*. To confirm that the presence of multiple cells expressing the *cdh-3*-driven AC reporter were due to defects in proliferation, we performed static imaging using a full-length translational *cdt-1*>CDT-1::GFP reporter that indicates cell cycle progression (Fig. 6A). As CDT-1 must be removed from origins of replication during DNA licensing, the transgene localizes to the nucleus during G1/G0 and is largely cytosolic at the onset of S phase (Matus et al., 2014; Matus et al., 2015). As expected, T444T (empty vector) and *fos-1(RNAi)* treated animals consistently exhibited nuclear localization of CDT-1::GFP (Fig. 6A; n $\geq$ 14 examined for each), indicating cell cycle arrest. However, following depletion of *egl-43, hlh-2* and *nhr-67* we identified multiple animals (Fig, 6A, n $\geq$ 10 for each RNAi) possessing non-invasive ACs lacking nuclear CDT-1::GFP, indicative of cycling ACs. Together, these results strongly support our molecular epistasis data suggesting that *egl-43* and *hlh-2* function upstream of *nhr-67* to maintain the post-mitotic state of the AC during invasion.

EGL-43 plays a role in cell cycle dependent and independent pro-invasive pathways

We have shown previously that the invasive activity of the AC following loss of NHR-67 can be completely rescued by restoring the AC to a post-mitotic G1/G0 state through upregulation of the cyclin-dependent kinase inhibitor, CKI-1 (p21/p27) (Matus et al., 2015). As our epistasis experiments revealed that *egl-43* and *hlh-2* positively co-regulate NHR-67 activity, we induced AC-specific expression of CKI-1 using a *cdh-3*>CKI-1::GFP integrated array and assessed invasion following RNAi depletions (Fig. 6B,C). As previously reported, induced localization of CKI-1::GFP in the AC completely rescued *nhr-67(RNAi)* treated animals (Fig. 6B,C; 100% invaded (n=42) as compared to 45% of control animals (n=62)). Additionally, *cdh-3*>CKI-1::GFP rescued AC invasion following depletion of *hlh-2* in L2 RNAi-feeding experiments in most animals examined (Fig. 6B,C; 87% invaded (n=77), as compared to 56% of control animals (n=59)). As expected, inducing G1/G0 arrest in the AC failed to rescue *fos-1(RNAi)* treated animals (Fig. 6B,C; 20% invaded (n=35), as compared to 51% of control animals (n=53)). Strikingly, although AC-specific CKI-1::GFP blocked AC proliferation following *egl-43(RNAi),* induced arrest failed to rescue invasion (Fig. 6B,C; 10% invaded (n=31), as compared to 19% of control animals (n=37)), suggesting that *egl-43* has pro-invasive functions outside of the cell cycle-dependent pathway. In support of a role for *egl-43* outside of cell cycle control, induced CKI-1::GFP restores the expression of a reporter of MMP activity (*zmp-1>mCherry*) in *nhr-67*-depleted animals, but neither *fos-1* nor *egl*-43-depleted animals show restoration of *zmp-1* transcription (Fig. 6D). Together, these results reveal the presence of two integrated sub-circuits necessary for invasive activity, with *egl-43* likely functioning in a critical role for both circuits.

EGL-43 isoforms function redundantly and in an autoregulative manner during AC invasion

Two functional isoforms of *egl-43* are encoded in the *C. elegans* genome (Fig. 7A). Previous research has suggested that the longer isoform functions downstream of *fos-*

342 *1a* to modulate MMP expression and other *fos-1* transcriptional targets, including *cdh-3*
343 and *him-4*/hemicentin (Hwang et al., 2007; Rimann and Hajnal, 2007). The shorter
344 isoform has been predicted to either function in Notch/Delta-mediated AC specification
345 (Hwang et al., 2007) and later Notch/Delta-mediated patterning of the ventral uterus
346 (Hwang et al., 2007) or as a potential competitive inhibitor for long isoform binding of
347 downstream targets (Rimann and Hajnal, 2007). Additionally, *fos-1a* and *egl-43* have
348 been shown to function in an incoherent feed-forward loop with negative feedback, with
349 *fos-1a* positively regulating and *egl-43* negatively regulating the levels of *mig-*
350 *10*/lamellipodin, a key adhesion protein important for stabilizing the attachment of the
351 AC to the BM (Wang et al., 2014). Titrating levels of MIG-10 is critical, as
352 overexpression of MIG-10 results in AC invasion defects (Wang et al., 2014). Although it
353 is readily apparent that levels of EGL-43 are critical, it is unclear from previous studies
354 whether the two isoforms have divergent functions during invasion. Thus, we next
355 decided to explore *egl-43* isoform function by using newly available tools, including
356 CRISP/Cas9-genome engineering and improved RNAi.
357
358 As the start codon methionine of the short isoform of *egl-43* is in-frame with and
359 contained within the long isoform, there was no simple way to solely tag or interfere with
360 the short isoform (Fig. 7A). However, we were able to generate a knock-in allele of GFP
361 at the N-terminus of *egl-43* to tag the long isoform specifically *(GFP::egl-43L*). This
362 allowed us to compare expression patterns of the long isoform to the internally GFP-
363 tagged allele that should dually label both isoforms (Fig. 7A,B). We examined animals
364 during uterine-vulval development and found overlapping expression patterns between
365 both isoforms with strong AC/VU/DU expression in the somatic gonad (Fig. 7B). Next,
366 we generated an *egl-43L*-specific improved RNAi targeting vector in T444T and
367 examined *egl-43L* depletion in the uterine-specific RNAi sensitive genetic background
368 (Hagedorn et al., 2009; Matus et al., 2010). Depletion of *egl-43L* resulted in a penetrant
369 AC invasion defect, as expected, but notably, we detected the presence of animals with
370 both a single non-invasive AC (19%) and animals with multiple ACs (39%) (Fig. 7C,D),
371 similar to depletion of both isoforms (Fig. 2). We then assessed if the long isoform has
372 the same quantitative relationship to the other TFs as observed when targeting both

isoforms. Indeed, *egl-43(RNAi)* and *egl-43L(RNAi)* exhibited comparable levels of depletion of *egl-43::GFP::egl-43* (75%, 78%), *GFP::hlh-2* (66%, 60%), and *nhr-67::GFP* (66%, 64%, respectively) ($n \geq 25$ animals examined per treatment) (Fig. 3E).

The insertion of GFP with a ~6 kb self-excising cassette (SEC) generates a transcriptional reporter that often interferes with the native transcriptional machinery at the locus where it is inserted. To generate an in-frame fusion protein with GFP, we excise the SEC by CRE-lox recombination via heat shock (Dickinson and Goldstein, 2016). To determine if newly generated alleles exhibit loss-of-function phenotypes and to examine transcriptional output, we next examined the pre-floxed versions of both edited knock-in alleles, in the presence of an mT1 balancer, as homozygous animals are non-fertile. Unfortunately, the allele resulting from the SEC insertion of the internal GFP-tag had an extremely low frequency of escapers, and we were unable to obtain L3-stage animals to examine for AC invasion defects. However, similar to our results with RNAi (Fig. 7C-E), an allele resulting from SEC insertion with the long isoform (*egl-43(bmd135)*) (Fig. 7F) displayed a 31% AC invasion defect (n=11/36), with animals containing either single or multiple non-invading ACs (Fig. 7G). Finally, as previous reports based on *promoter>GFP* transgenic animals have suggested that *egl-43* may be autoregulatory (Matus et al., 2010; Rimann and Hajnal, 2007; Wang et al., 2014), we examined the mT1-balanced pre-floxed allele of the internally-tagged *egl-43* targeting both isoforms (*egl-43::GFP^SEC^::egl-43*). We found strong evidence of autoregulation as *egl-43(RNAi)* reduced the expression of GFP by 65% ($n \geq 25$ animals) (Fig. 7H,I). Taken together, these results suggest that both isoforms of *egl-43* function redundantly to regulate multiple transcriptional sub-circuits critical for the establishment of the invasive phenotype (Fig. 8).

DISCUSSION

Cellular invasion requires the coordination of extrinsic cues from the surrounding microenvironment and orchestration of intrinsic regulatory circuits (Rowe and Weiss, 2008; Sherwood and Plastino, 2018). We focus here on exploring the intrinsic pathways

that autonomously regulate *C. elegans* AC invasion. This study was tractable due to several recent technological breakthroughs, allowing for examination of endogenous regulatory relationships quantitatively. First, we created GFP knock-in alleles of the four transcription factors (*fos-1a*, *egl-43*, *hlh-2* and *nhr-67*) previously implicated in autonomously regulating AC invasion (Hwang et al., 2007; Matus et al., 2015; Rimann and Hajnal, 2007; Schindler and Sherwood, 2011; Sherwood et al., 2005). Additionally, we generated new improved RNAi depletion constructs to strongly reduce TF activity. Combining these technologies enabled us to correlate phenotype to protein depletion and compare regulatory relationships in a quantitative framework. Our experiments reveal new roles for *egl-43* and *hlh-2* in regulating *nhr-67*-mediated cell cycle arrest and suggest that these TFs function in a coherent feed-forward loop with positive feedback. Additionally, we demonstrate that *egl-43* is autoregulatory and functions in both cell cycle-dependent and independent sub-circuits to orchestrate invasion (Fig. 8). Together, these results provide the first description of the regulatory relationships between the endogenous TFs that promote invasive differentiation during *C. elegans* uterine-vulval attachment.

nhr-67 functions to maintain the post-mitotic state of the AC (Matus et al., 2015). Here, we demonstrate that *egl-43* and *hlh-2* both regulate *nhr-67*, and that *nhr-67*-depletion also reduces HLH-2 levels. Network inference suggests these three TFs function in a classic type I coherent feed-forward loop regulatory circuit with positive feedback (Mangan and Alon, 2003) (Fig. 8). Our model fits with data by others showing that *hlh-2* may directly bind to the *nhr-67* promoter through two canonical E-box motifs found in a 164 bp window in a functional promoter element deleted in several hypomorphic alleles of *nhr-67* (Bodofsky et al., 2018; Matus et al., 2015; Verghese et al., 2011). Further, deletion of these *hlh-2* conserved binding sites results in loss of AC-specific GFP expression in transgenic lines (Bodofsky et al., 2018). While others have also reported positive regulation by EGL-43 on *nhr-67* activity via transgenic reporters (Bodofsky et al., 2018), this interaction may be indirect, as no known binding sites exist for EGL-43 in the *nhr-67* promoter. Alternatively, EGL-43 and NHR-67 have been predicted, based on

yeast two-hybrid experiments, to interact at the protein-protein level (Reece-Hoyes et al., 2013).

We have also examined the activity of the two previously identified isoforms of *egl-43* (Hwang et al., 2007; Rimann and Hajnal, 2007), by generating multiple GFP knock-in alleles paired with more penetrant RNAi depletions. Previous work has suggested that the long isoform functions specifically downstream of *fos-1a* to regulate MMP activity and other pro-invasive gene expression (Rimann and Hajnal, 2007). As both mutant analyses and RNAi depletion of the long isoform phenocopy depletion of both isoforms, our data suggest that either the long isoform of *egl-43* functions redundantly with the short isoform or that it functions as the dominant isoform at the intersection of multiple regulatory circuits.

Together, our data and corroborating evidence from the literature support the existence of a coherent feed-forward circuit with positive feedback among *egl-43*, *hlh-2* and *nhr-67,* in maintaining the post-mitotic state of the AC (Fig. 8). We have previously shown that *nhr-67* positively regulates transcripts of *cki-1* specifically in the AC (Matus et al., 2015). Thus, to test our putative regulatory circuit, we induced expression of *cki-1* in the AC, which prevents an AC from inappropriately entering the cell cycle. This forced G1/G0 arrest fully rescues *nhr-67* depletion (Matus et al., 2015). Strikingly, induced *cki-1* strongly rescues *hlh-2* depletion as well. Thus, *hlh-2* appears to function primarily in controlling NHR-67 activity in maintenance of the post-mitotic state of the AC. These results are in stark contrast to *cki-1* induction paired with strong *egl-43* depletion, where most cell cycle arrested ACs still fail to invade, and fail to express an MMP, *zmp-1*. These results suggest that *egl-43* functions in multiple regulatory circuits. This is supported by recent work demonstrating a type I incoherent feed-forward loop between *fos-1a, egl-43L* and the BM adhesion protein MIG-10/lamellipodin (Wang et al., 2014). Finally, our observation that *egl-43* is autoregulatory, also shown previously by *promoter>GFP* fusions (Matus et al., 2015; Rimann and Hajnal, 2007; Wang et al., 2014) supports a model where *egl-43* occupies a key node at the intersection of multiple pro-invasive circuits (Fig. 8). As all three TFs that function in the cell cycle-dependent

circuit are also involved in the Notch/Delta-mediated specification of the AC/VU fates (Hwang et al., 2007; Karp and Greenwald, 2004) it will be interesting to explore whether these regulatory relationships are maintained in the AC/VU specification network.

Feed-forward regulatory loops are likely the most well described network motif occurring in all GRNs (Cordero and Hogeweg, 2006; Davidson, 2010; Mangan and Alon, 2003). From *E. coli* (Milo et al., 2002; Shen-Orr et al., 2002) and yeast (Lee et al., 2002; Milo et al., 2002) to a myriad of examples across the Metazoa (reviewed in Davidson, 2010), feed-forward loops are thought to function as filters for transient inputs (reviewed in Alon, 2007; Hinman, 2016). The addition of positive feedback, generating a coherent feed-forward loop, provides stability to the sub-circuit and is often found in differentiation gene batteries coincident with autoregulation (Davidson, 2010). These network motifs have been described in many developmental contexts, from MyoD-driven vertebrate skeletal muscle differentiation (Penn et al., 2004), the patterning of the Drosophila egg shell via the TF Broad and interactions with EGFR and Dpp signaling (Yakoby et al., 2007), Pax6-dependent regulation of c-Maf and crystallin expression in the mouse embryonic lens (Xie and Cvekl, 2011) and terminal selector neuronal differentiation in *C. elegans* and mammals (reviewed in Hobert, 2008). During *C. elegans* embryonic development, a series of coherent feed-forward loops utilizing SKN-1/MED-1,2 and then a suite of reiteratively used GATA factors (END-1,-3) are required to successfully pattern endomesoderm development (reviewed in Maduro, 2009). Thus, our identification of a coherent feed-forward loop in the *C. elegans* AC maintaining the post-mitotic state, likely evolved to coordinate AC cell cycle exit, both as a by-product of terminal differentiation and a morphogenetic requirement for invasive behavior (Matus et al., 2015). AC invasion is necessary for egg-laying, and defects in the process results in penetrant Protruding vulva/Egg-laying defective (Pvl/Egl) phenotypes, reducing fecundity in individual animals by nearly ten-fold. Thus, redundant control of the sub-circuit regulating differentiation and cell cycle arrest may have been under strong selection, providing an explanation for the regulatory relationships between *egl-43*, *hlh-2* and *nhr-67* we characterize here.

We are just beginning to understand connections between regulatory circuits and morphogenetic behaviors (Martik and McClay, 2015; Saunders and McClay, 2014) (Christiaen et al., 2008). It is our hope that the ease of genome editing and protein perturbation strategies will facilitate the kind of analyses we have performed here in other metazoan systems. In summary, in this study we characterize the complex relationships between the four pro-invasive TFs that function during AC differentiation to program invasive behavior. We identify a classic type I FFL regulating mitotic exit and controlling a switch between proliferative and invasive behavior. Whether or not similar circuit architecture is utilized to regulate invasive and proliferative cell biology in other developmental invasive contexts, including mammalian trophoblast implantation and placentation (Carter et al., 2015; Red-Horse et al., 2004) and EMT events during gastrulation (Vega et al., 2004) is still poorly understood. Finally, as there appear to be many cancer sub-types that may switch between proliferative and invasive fates (reviewed in Kohrman and Matus, 2017), improving our understanding of the transcriptional network architecture of invasive cells may provide new therapeutic nodes to target in reducing the lethality associated with cancer metastasis.

MATERIALS AND METHODS

C. elegans strains and culture conditions

Animals were reared under standard conditions and cultured at 25°C, with the exception of temperature-sensitive strains containing the *rrf-3(pk1426)* allele, conferring RNAi hypersensitivity, which were maintained at 15°C and 20°C (Brenner, 1974; Simmer et al., 2002). Animals were synchronized for experiments through alkaline hypochlorite treatment of gravid adults to isolate eggs (Porta-de-la-Riva et al., 2012). In the text and figures, we designate linkage to a promoter through the use of the (>) symbol and fusion of a proteins via a (::) annotation. The following transgenes and alleles were used in this study: *qyIs102[fos-1>RDE-1;myo-2>YFP]* **LG I** *hlh-2(bmd90[hlh-2>LoxP::GFP::HLH-2]), qyIs227 [cdh-3>mCherry::moeABD];* **LG II** *egl-43(bmd87[egl-43>SEC::GFP::EGL-43]), egl-43(bmd88[egl-43>LoxP::GFP::EGL-43]), egl-43(bmd136[egl-43L>LoxP::GFP::EGL-*

43]) rrf-3(pk1426); qyIs17 [zmp-1>mCherry] **LG III** *unc-119(ed4),* **LG IV** *nhr-67(syb509[nhr-67>NHR-67::GFP]), qyIs10[lam-1>LAM-1::GFP]* **LG V** *fos-1(bmd138[fos-1>LoxP::GFP::FOS-1]), qyIs225[cdh-3>mCherry::moeABD], rde-1(ne219), qyIs24[cdh-3$^{1.5}$>mCherry::PLCδPH], qyIs266[cdh-3>CKI-1::GFP]* **LG X** *qyIs7[lam-1>LAM-1::GFP].* See Table S2 for additional details of strains used and generated in this study.

Molecular biology and microinjection

Transcription factors were tagged at their endogenous loci using CRISPR/Cas9 genome editing via microinjection into the hermaphrodite gonad (Dickinson and Goldstein, 2016; Dickinson et al., 2013). Repair templates were generated as synthetic DNAs from either Integrated DNA Technologies (IDT) as gBlocks or Twist Biosciences and cloned into *ccdB* compatible sites in pDD282 by New England Biolabs Gibson assembly (Dickinson et al. 2016). Homology arms ranged from 690 – 1200 bp (see supplementary material Table S3-5 for additional details). sgRNAs were constructed by EcoRV and NheI digestion of plasmid pDD122. A 230bp amplicon was generated replacing the sgRNA targeting sequence from pDD122 with a new sgRNA and NEB Gibson assembly was used to generate new sgRNA plasmids (see Table S3,4 for more details). Hermaphrodite adults were co-injected with guide plasmid (50 ng/μL), repair plasmid (50 ng/μL), and an extrachromosomal array marker (pCFJ90, 2.5 ng/μL), and incubated at 25 ºC for several days before carrying out screening and floxing methods associated with the SEC (Dickinson et al., 2015).

RNA interference

An RNAi library of the pro-invasive TFs was constructed by cloning 950-1000 bp of synthetic DNA (663 bp for the *egl-43* long-specific isoform) based on cDNA sequences available on WormBase (www.wormbase.org) into the highly efficient T444T RNAi vector (Grove et al., 2018; Sturm et al., 2018). Synthetic DNAs were generated by IDT as gBlocks or Twist Biosciences as DNA fragments and cloned into restriction digested T444T using NEB Gibson Assembly (see Tables S3, S6 for additional details). For most

558 experiments, synchronized L1 stage animals were directly exposed to RNAi through
559 feeding with bacteria expressing dsRNA (Conte Jr. et al., 2015). Due to the fact that
560 early *hlh-2* RNAi treatment perturbs AC specification, animals were initially placed on
561 empty vector RNAi plates and then transferred to *hlh-2* RNAi plates post-specification,
562 approximately 12 hours later at 25°C or 24 hours later at 15°C (Schindler and
563 Sherwood, 2011).
564
565 ***Live cell imaging***
566
567 All micrographs included in this manuscript were collected on a Hamamatsu Orca EM-
568 CCD camera mounted on an upright Zeiss AxioImager A2 with a Borealis-modified
569 CSU10 Yokagawa spinning disk scan head using 488 nm and 561 nm Vortran lasers in
570 a VersaLase merge and a Plan-Apochromat 100x/1.4 (NA) Oil DIC objective.
571 MetaMorph software (Molecular Devices) was used for microscopy automation. Several
572 experiments were imaged using epifluorescence collected on a Zeiss Axiocam MRM
573 camera, also mounted on an upright Zeiss AxioImager A2 and a Plan-Apochromat
574 100x/1.4 (NA) Oil DIC objective. Animals were mounted into a drop of M9 on a 5%
575 Noble agar pad containing approximately 10 mM sodium azide as paralytic and topped
576 with a coverslip.
577
578 ***Scoring of AC invasion***
579
580 AC invasion was scored at the P6.p 4-cell stage, when 100% of wild-type animals
581 exhibit a breached BM (Sherwood and Sternberg, 2003). Presence of green
582 fluorescence under the AC, in strains with the laminin::GFP transgene, or presence of a
583 phase dense line, in strains without the transgene, was used to assess invasion. Wild-
584 type invasion is defined as a breach as wide as the basolateral surface of the AC,
585 whereas partial invasion indicates the presence of a breach smaller than the footprint of
586 the AC (Sherwood and Sternberg, 2003).
587
588 ***Image quantification and statistical analyses***

589

590 Images were processed using FIJI/ImageJ (v. 2.0.0-rc-69, NIH Image) (Schindelin et al.,
591 2012). With the exception of the developmental series displayed in Fig. 3 & Fig. S1D, all
592 other micrographs were taken at the P6.p 4-cell stage, when 100% of wild-type animals
593 exhibit a breached BM. Expression levels of TFs were quantified by measuring the
594 mean gray value of AC nuclei of, defined as somatic gonad cells strongly expressing the
595 *cdh-3>mCherry::moeABD* transgene, subtracted by the mean gray value of a
596 background region of equal area, to account for EM-CCD camera noise. *zmp-*
597 *1>mCherry* levels were quantified in control and *cdh-3>CKI-1*::GFP animals by
598 measuring the mean gray value of the entire AC, selected either by a hand drawn
599 region of interest or using the threshold and wand tool in FIJI/ImageJ (v. 1.52n). For
600 molecular epistasis experiments (Fig. 5, Fig. 7E,I & Fig. S3) and characterization of
601 CDT-1 localization (Fig. 6), only TF RNAi treated animals that did not exhibit defects in
602 invasion were excluded from analysis. Data was normalized to negative control (empty
603 vector) values for the plots in Fig. 4E, 6D, & 7E, and to both negative control and
604 positive control values for determining the interaction strengths represented in Fig. 8.
605 Images were overlaid and figures were assembled using Adobe Photoshop and
606 Illustrator CS, respectively. Statistical analyses and plotting of data were conducted
607 using RStudio (v. 1.1.463). Individual data points for each experiment were collected
608 over multiple days. Statistical significance was determined using either a two-tailed
609 Student's t-test or Fisher's exact probability test. Figure legends specify when each test
610 was used and the p-value cut-off set.

611

612 **ACKNOWLEDGMENTS**

613 We are grateful to R. Adikes, B. Kinney, A. Kohrman, M. Martinez and B. Martin for
614 comments on the manuscript.

615

616 **COMPETING INTERESTS**

617 No competing interests declared.

618

619 **AUTHOR CONTRIBUTIONS**

D.Q.M., T.N.M. and J.J.S. designed the experiments. All authors performed the experiments. T.N.M. and D.Q.M. performed the data analyses and wrote the paper with feedback from the other authors.

FUNDING

This work was funded by the National Institute of Health (NIH) National Cancer Institute [5R00CA154870-05 to D.Q.M.] and National Institute of General Medical Sciences (NIGMS) [1R01GM121597-01 to D.Q.M.]. D.Q.M. is also a Damon Runyon-Rachleff Innovator supported (in part) by the Damon Runyon Cancer Research Foundation [DRR-47-17]. J.J.S. is supported by NIGMS [3R01GM121597-02S1] and N.J.P. is supported by the American Cancer Society [132969-PF-18-226-01-CSM]. Some strains were provided by the *Caenorhabditis* Genetics Center, which is funded by the NIH Office of Research Infrastructure Programs [P40 OD010440].

REFERENCES

Alon, U. (2007). Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* **8**, 450–61.

Bodofsky, S., Liberatore, K., Pioppo, L., Lapadula, D., Thompson, L., Birnbaum, S., McClung, G., Kartik, A., Clever, S. and Wightman, B. (2018). A tissue-specific enhancer of the C. elegans nhr-67/tailless gene drives coordinated expression in uterine stem cells and the differentiated anchor cell. *Gene Expr. Patterns* **30**, 71–81.

Brenner, S. (1974). The genetics of *Caenorabditis elegans*. *Genetics* **77**, 71–94.

Carter, A. M., Enders, A. C. and Pijnenborg, R. (2015). The role of invasive trophoblast in implantation and placentation of primates. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20140070–20140070.

Christiaen, L., Davidson, B., Kawashima, T., Powell, W., Nolla, H., Vranizan, K. and Levine, M. (2008). The transcription/migration interface in heart precursors of Ciona intestinalis. *Science (80-.).* **320**, 1349–1352.

Conte Jr., D., T. Macneil, L., J. M. Walhout, A. and C. Mello, C. (2015). *RNA Interference in Caenorhabditis Elegans*.

Cordero, O. X. and Hogeweg, P. (2006). Feed-forward loop circuits as a side effect of genome evolution. *Mol. Biol. Evol.* **23**, 1931–1936.

Davidson, E. H. (2010). Emerging properties of animal gene regulatory networks. *Nature* **468**, 911–920.

Dickinson, D. J. and Goldstein, B. (2016). CRISPR-based methods for caenorhabditis elegans genome engineering. *Genetics* **202**, 885–901.

Dickinson, D. J., Ward, J. D., Reiner, D. J. and Goldstein, B. (2013). Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination. *Nat. Methods* **10**, 1028–1034.

Dickinson, D. J., Pani, A. M., Heppert, J. K., Higgins, C. D. and Goldstein, B. (2015). Streamlined genome engineering with a self-excising drug selection cassette. *Genetics* **200**, 1035–1049.

Grove, C., Cain, S., Chen, W. J., Davis, P., Harris, T., Howe, K., Kishore, R., Lee, R.,

Paulini, M., Raciti, D., et al. (2018). Using WormBase, a genome biology resource for Caenorhabditis elegans and related nematodes. *Methods Mol. Biol.* **1757**, 399–470.

Gupta, B. P., Hanna-Rose, W. and Sternberg, P. W. (2012). Morphogenesis of the vulva and the vulval-uterine connection. *WormBook*.

Haerty, W., Artieri, C., Khezri, N., Singh, R. S. and Gupta, B. P. (2008). Comparative analysis of function and interaction of transcription factors in nematodes: Extensive conservation of orthology coupled to rapid sequence evolution. *BMC Genomics* **9**, 1–16.

Hagedorn, E. J., Yashiro, H., Ziel, J. W., Ihara, S., Wang, Z. and Sherwood, D. R. (2009). Integrin Acts Upstream of Netrin Signaling to Regulate Formation of the Anchor Cell's Invasive Membrane in C. elegans. *Dev. Cell* **17**, 187–198.

Hagedorn, E. J., Ziel, J. W., Morrissey, M. A., Linden, L. M., Wang, Z., Chi, Q., Johnson, S. A. and Sherwood, D. R. (2013). The netrin receptor DCC focuses invadopodia-driven basement membrane transmigration in vivo. *J. Cell Biol.* **201**, 903–913.

Hanahan, D. and Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell* **144**, 646–674.

Hinman, V. F. (2016). Gene Networks Driving Development, Conservation and Evolution of. *Encycl. Evol. Biol.* 110–116.

Hobert, O. (2008). Regulatory logic of neuronal diversity: Terminal selector genes and selector motifs. *Proc. Natl. Acad. Sci.* **105**, 20067–20071.

Hwang, B. J., Meruelo, A. D. and Sternberg, P. W. (2007). C. elegans EVI1 proto-oncogene, EGL-43, is necessary for Notch-mediated cell fate specification and regulates cell invasion. *Development* **134**, 669–679.

Karp, X. and Greenwald, I. (2004). Multiple roles for the E/Daughterless ortholog HLH-2 during C. elegans gonadogenesis. *Dev. Biol.* **272**, 460–469.

Kelley, L. C., Lohmer, L. L., Hagedorn, E. J. and Sherwood, D. R. (2014). Traversing the basement membrane in vivo: A diversity of strategies. *J. Cell Biol.* **204**, 291–302.

Kelley, L. C., Hastie, E. L., Cáceres, R., Chi, Q., Schindler, A. J., Jiang, Y., Matus,

D. Q., Plastino, J. and Sher (2018). Adaptive F-actin polymerization and localized ATP production drive basement membrane invasion in the absence of MMPs. *Dev. Cell* 1–16.

Kohrman, A. Q. and Matus, D. Q. (2017). Divide or Conquer: Cell Cycle Regulation of Invasive Behavior. *Trends Cell Biol.* **27**, 12–25.

Lee, T., Hannett, N., Harbison, C., Thompson, C., Simon, I., Zeitlinger, J., Jennings, E., Murray, H., Gordon, D., Ren, B., et al. (2002). Transcriptional Regulatory Networks in Saccharomyces cerevisiae. *Science (80-.).* **298**, 799–804.

Maduro, M. F. (2009). Structure and evolution of the C. elegans embryonic endomesoderm network. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1789**, 250–260.

Mangan, S. and Alon, U. (2003). Structure and function of the feed-forward loop network motif. *Proc. Natl. Acad. Sci.* **100**,.

Martik, M. L. and McClay, D. R. (2015). Deployment of a retinal determination gene network drives directed cell migration in the sea urchin embryo. *Elife* **4**, 1–19.

Matus, D. Q., Li, X.-Y., Durbin, S., Agarwal, D., Chi, Q., Weiss, S. J. and Sherwood, D. R. (2010). In vivo identification of regulators of cell invasion across basement membranes. *Sci. Signal.* **3**, ra35.

Matus, D. Q., Chang, E., Makohon-Moore, S. C., Hagedorn, M. A., Chi, Q. and Sherwood, D. R. (2014). Cell division and targeted cell cycle arrest opens and stabilizes basement membrane gaps. *Nat. Commun.* **5**,.

Matus, D. Q., Lohmer, L. L., Kelley, L. C., Schindler, A. J., Kohrman, A. Q., Barkoulas, M., Zhang, W., Chi, Q. and Sherwood, D. R. (2015). Invasive Cell Fate Requires G1 Cell-Cycle Arrest and Histone Deacetylase-Mediated Changes in Gene Expression. *Dev. Cell* **35**, 162–174.

Medwig, T. N. and Matus, D. Q. (2017). Breaking down barriers: the evolution of cell invasion. *Curr. Opin. Genet. Dev.* **47**, 33–40.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D. and Alon, U. (2002). Network motifs: simple building blocks of complex networks. *Science (80-.).* **298**, 824–827.

Morrissey, M. A., Keeley, D. P., Hagedorn, E. J., McClatchey, S. T. H., Chi, Q., Hall,

D. H. and Sherwood, D. R. (2014). B-LINK: A Hemicentin, Plakin, and Integrin-Dependent Adhesion System that Links Tissues by Connecting Adjacent Basement Membranes. *Dev. Cell* **31**, 319–331.

Penn, B. H., Bergstrom, D. A., Dilworth, F. J., Bengal, E. and Tapscott, S. J. (2004). A MyoD -generated feed-forward circuit temporally patterns gene expression dining skeletal muscle differentiation. *Genes Dev.* **18**, 2348–2353.

Porta-de-la-Riva, M., Fontrodona, L., Villanueva, A. and Cerón, J. (2012). Basic Caenorhabditis elegans Methods: Synchronization and Observation. *J. Vis. Exp.* 1–9.

Red-Horse, K., Zhou, Y., Genbacev, O., Prakobphol, A., Foulk, R., McMaster, M. and Fisher, S. J. (2004). Trophoblast differentiation during embryo implantation and formation of the maternal-fetal interface. *J. Clin. Invest.* **114**, 744–754.

Reece-Hoyes, J. S., Pons, C., Diallo, A., Mori, A., Shrestha, S., Kadreppa, S., Nelson, J., DiPrima, S., Dricot, A., Lajoie, B. R., et al. (2013). Extensive rewiring and complex evolutionary dynamics in a C.elegans multiparameter transcription factor network. *Mol. Cell* **51**, 116–127.

Rimann, I. and Hajnal, A. (2007). Regulation of anchor cell invasion and uterine cell fates by the egl-43 Evi-1 proto-oncogene in Caenorhabditis elegans. *Dev. Biol.* **308**, 187–195.

Rowe, R. G. and Weiss, S. J. (2008). Breaching the basement membrane: who, when and how? *Trends Cell Biol.* **18**, 560–574.

Sallee, M. D., Littleford, H. E. and Greenwald, I. (2017). A bHLH Code for Sexually Dimorphic Form and Function of the C. elegans Somatic Gonad. *Curr. Biol.* **27**, 1853-1860.e5.

Saunders, L. R. and McClay, D. R. (2014). Sub-circuits of a gene regulatory network control a developmental epithelial-mesenchymal transition. *Development* **141**, 1503–1513.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–82.

Schindler, A. J. and Sherwood, D. R. (2011). The transcription factor HLH-

2/E/Daughterless regulates anchor cell invasion across basement membrane in C. elegans. *Dev. Biol.* **357**, 380–391.

Shen-Orr, S. S., Milo, R., Mangan, S. and Alon, U. (2002). Network motifs in the transcriptional regulation network of Escherichia coli. *Nat. Genet.* **31**, 64–68.

Sherwood, D. R. and Plastino, J. (2018). Invading, leading and navigating cells in caenorhabditis elegans: Insights into cell movement in vivo. *Genetics* **208**, 53–78.

Sherwood, D. R. and Sternberg, P. W. (2003). Anchor cell invasion into the vulval epithelium in C. elegans. *Dev. Cell* **5**, 21–31.

Sherwood, D. R., Butler, J. A., Kramer, J. M. and Sternberg, P. W. (2005). FOS-1 promotes basement-membrane removal during anchor-cell invasion in C. elegans. *Cell* **121**, 951–962.

Simmer, F., Tijsterman, M., Parrish, S., Koushika, S. P., Nonet, M. L., Fire, A., Ahringer, J., Plasterk, R. H. A., Louis, S., Road, T. C., et al. (2002). Loss of the Putative RNA-Directed RNA Polymerase RRF-3 Makes. *Current* **12**, 1317–1319.

Sturm, Á., Saskoi, É., Tibor, K., Weinhardt, N. and Vellai, T. (2018). Highly efficient RNAi and Cas9-based auto-cloning systems for C. elegans research. *Nucleic Acids Res.* **46**, e105.

Vega, S., Morales, A. V., Ocaña, O. H., Valdés, F., Fabregat, I. and Nieto, M. A. (2004). Snail blocks the cell cycle and confers resistance to cell death. *Genes Dev.* **18**, 1131–1143.

Verghese, E., Schocken, J., Jacob, S., Wimer, A. M., Royce, R., Nesmith, J. E., Baer, G. M., Clever, S., McCain, E., Lakowski, B., et al. (2011). The tailless ortholog nhr-67 functions in the development of the C. elegans ventral uterus. *Dev. Biol.* **356**, 516–528.

Wang, L., Shen, W., Lei, S., Matus, D., Sherwood, D. and Wang, Z. (2014). MIG-10 (Lamellipodin) stabilizes invading cell adhesion to basement membrane and is a negative transcriptional target of EGL-43 in C. elegans. *Biochem. Biophys. Res. Commun.* **452**, 328–333.

Wilkinson, H. A., Fitzgerald, K. and Greenwald, I. (1994). Reciprocal changes in expression of the receptor lin-12 and its ligand lag-2 prior to commitment in a C. elegans cell fate decision. *Cell* **79**, 1187–1198.

Xie, Q. and Cvekl, A. (2011). The orchestration of mammalian tissue morphogenesis through a series of coherent feed-forward loops. *J. Biol. Chem.* **286**, 43259–43271.

Yakoby, N., Lembong, J., Schupbach, T. and Shvartsman, S. Y. (2007). Drosophila eggshell is patterned by sequential action of feedforward and feedback loops. *Development* **135**, 343–351.

FIGURES AND FIGURE LEGENDS

Figure 1 Medwig-Kinney et al., 2019

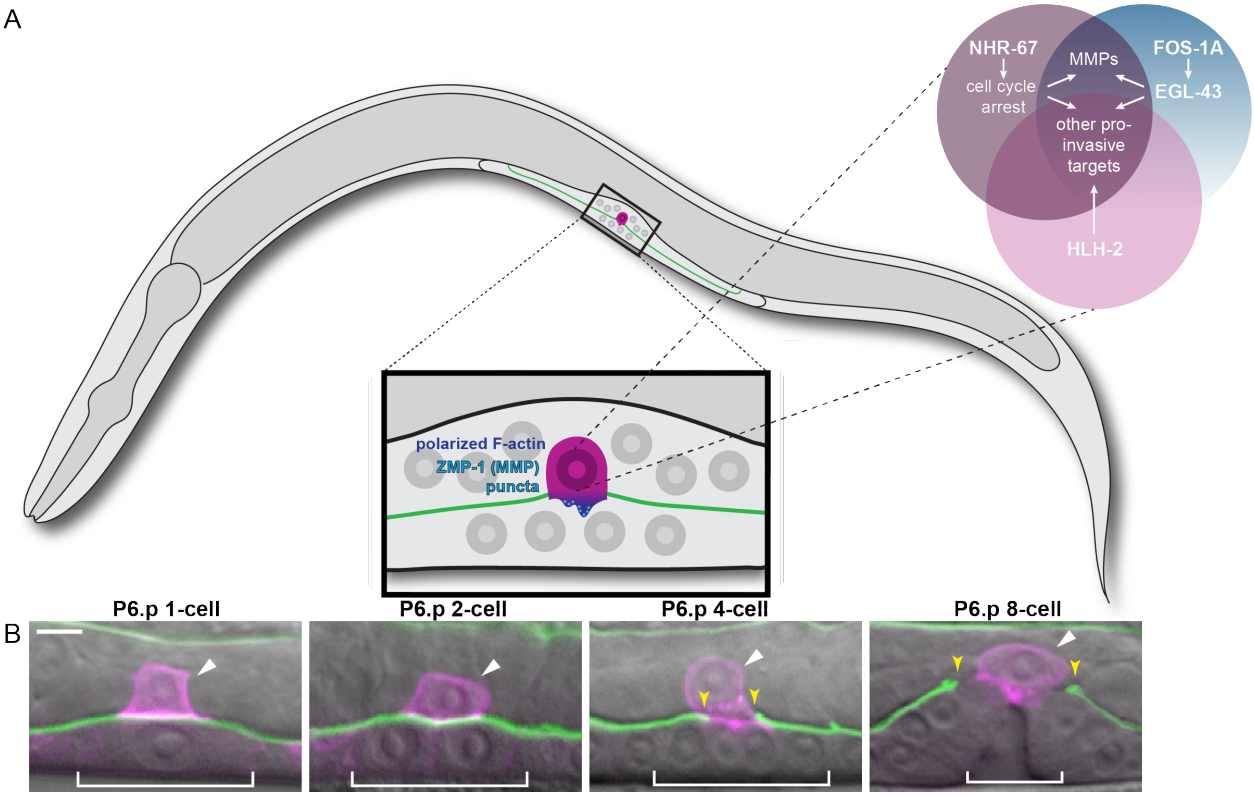


Fig. 1. Overview of *C. elegans* AC invasion. (A) Schematic depicting *C. elegans* AC (magenta) invasion through the BM (green), including polarization of the F-actin cytoskeleton (blue) into invasive protrusions and expression of MMPs (cyan) in the post-mitotic AC. Venn diagram (upper right) summarizes current TF relationships. (B) Single confocal planes overlaid on DIC depict the timing of AC (magenta, expressing *cdh-3*$^{1.5}$>mCherry::PLCδ^{PH}) invasion through the BM (green, *lam-1*>LAM-1::GFP). In this and all other figures, white arrowheads denote ACs, yellow arrowheads demarcate boundaries of gaps in the BM, and white brackets delineate P6.p and its descendants. Scale bar 5 μm, in this and all other figures.

Figure 2 Medwig-Kinney et al., 2019

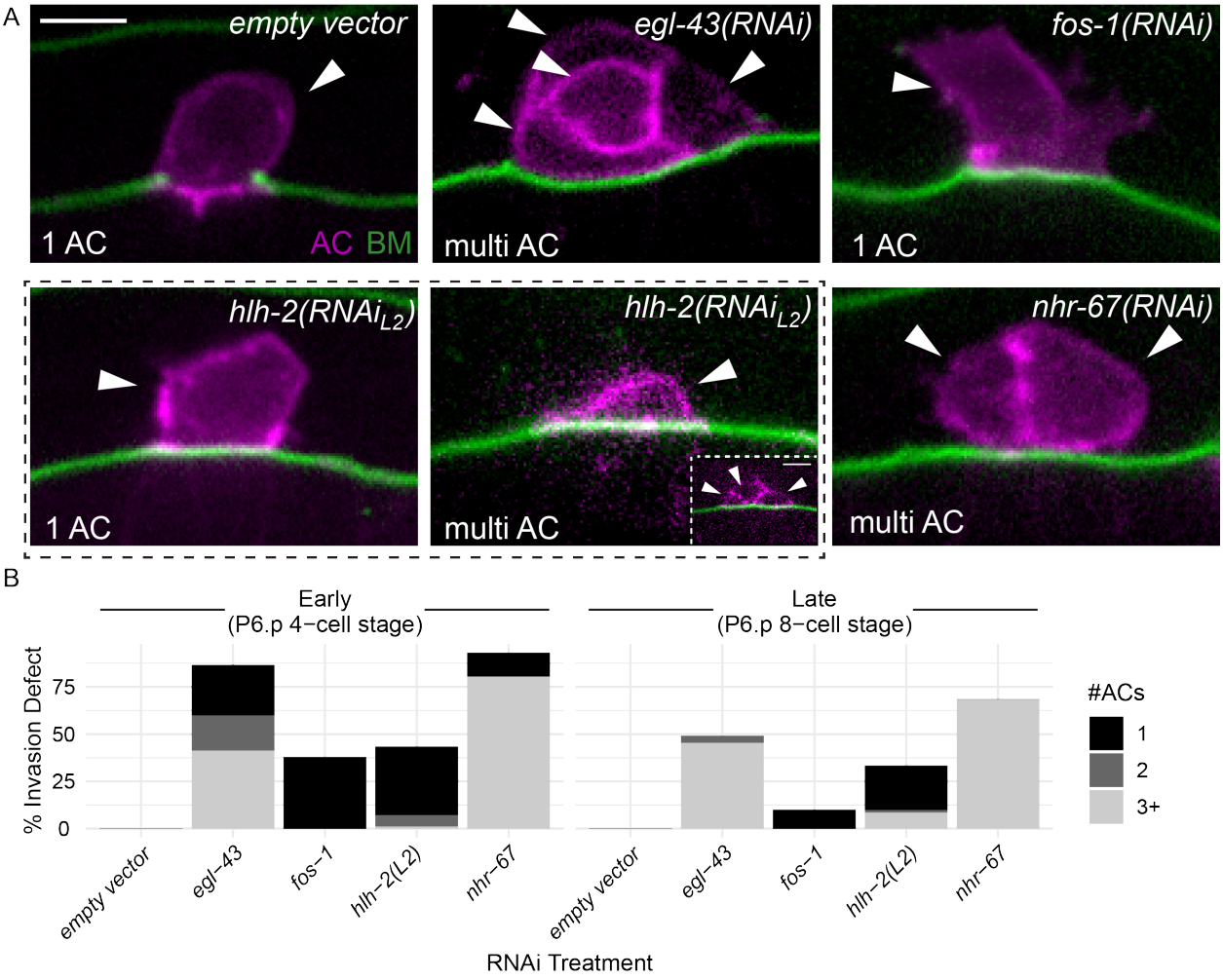


Fig. 2. RNAi depletion of pro-invasive TFs leads to defects in AC invasion and proliferation. (A) Single plane of confocal z-stack depicting representative empty vector control (upper left) and TF-RNAi depletion phenotypes. The AC-specific membrane marker (magenta, *cdh-3*$^{1.5}$>mCherry::PLCδ^{PH}) and BM marker (green, *lam-1*>LAM-1::GFP) are overlaid in each micrograph. In this and all other figures, dashed box denotes single vs. multi AC phenotypes that result from HLH-2 depletion, with multiple ACs from additional confocal z-planes shown as insets, accordingly. (B) Stacked bar chart showing penetrance of invasion defects, binned by number of ACs (p-value < 0.00001, Fisher's exact test, empty vector as compared to each TF-RNAi, n ≥ 50 animals examined for each treatment).

Figure 3 Medwig-Kinney et al., 2019

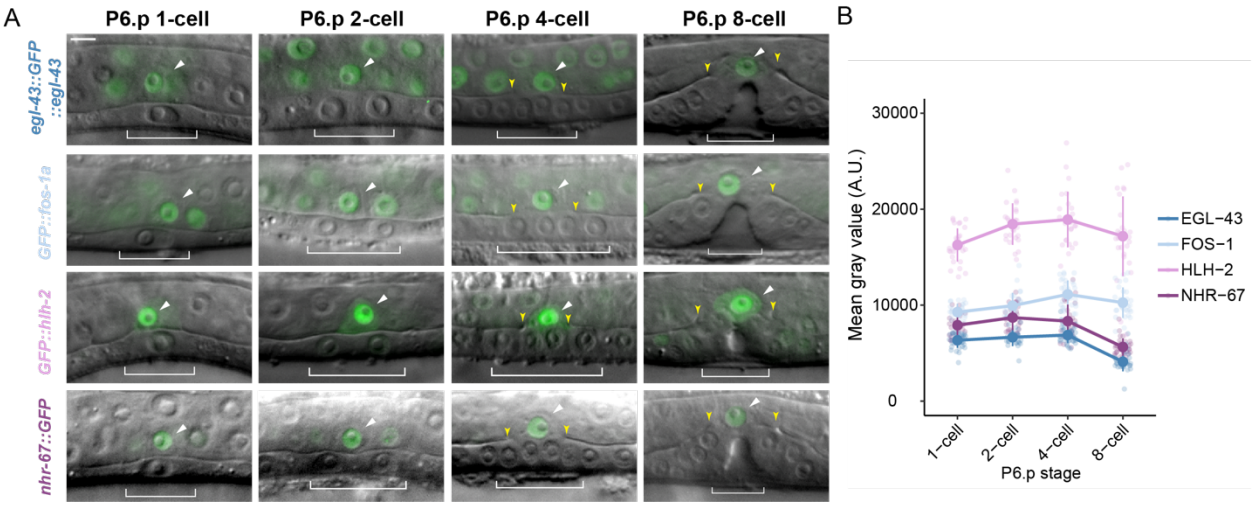
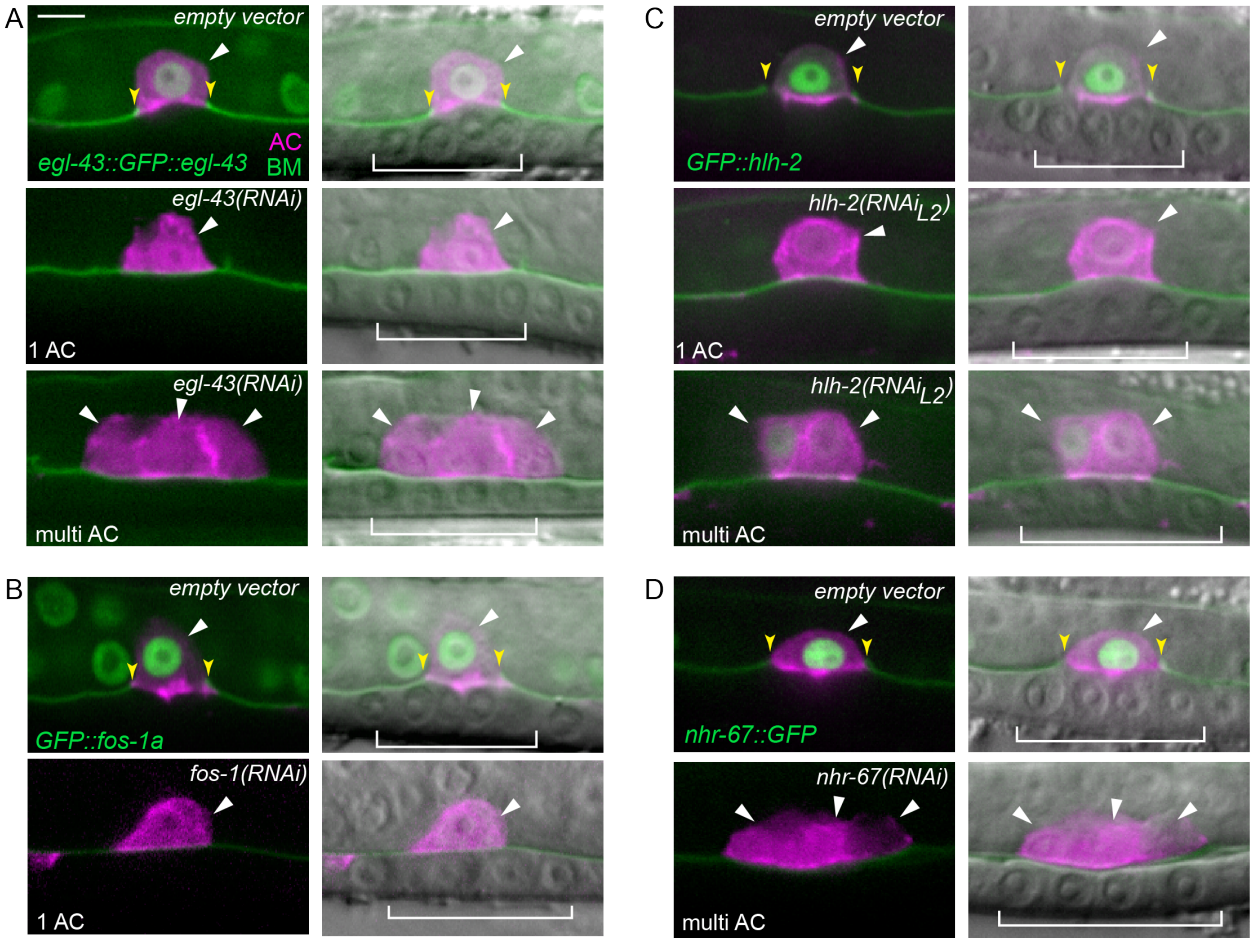


Fig. 3. Expression of endogenously-tagged pro-invasive TFs in the AC over developmental time. (A) DIC overlaid with single plane of confocal z-stack for endogenous GFP-tagged TFs. (B) Quantification of TF GFP, in relation to the division of P6.p. In this and all other figures, circle denotes mean gray value of population (n ≥ 25 for each stage), in this and all other figures bar denotes S.D.

Figure 4 Medwig-Kinney et al., 2019



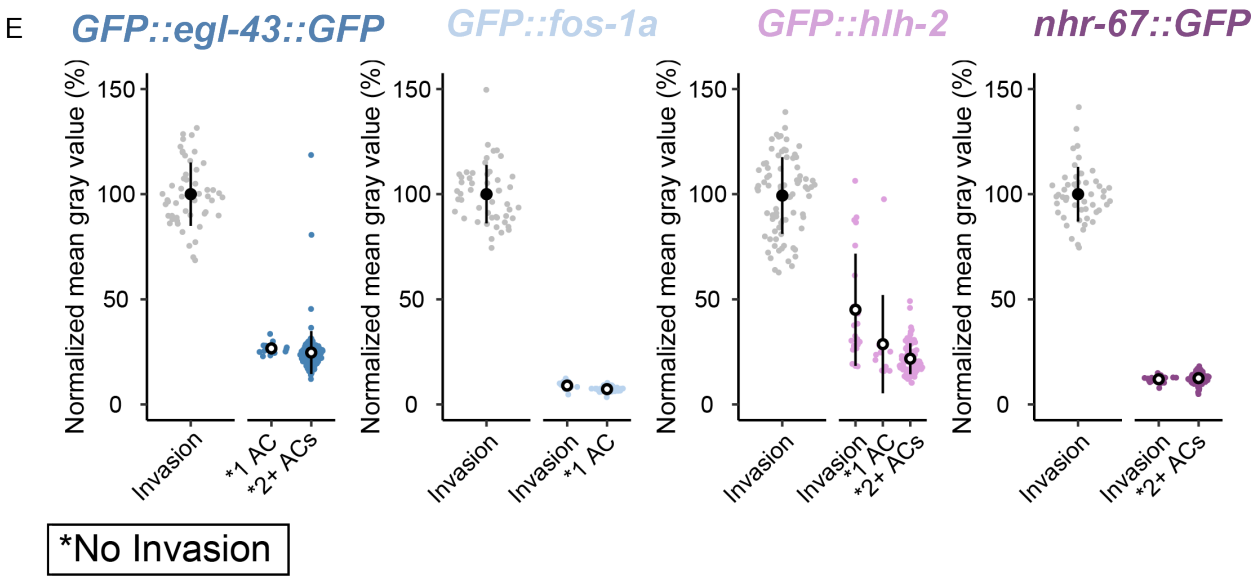


Fig. 4. Quantification of TF-RNAi phenotype penetrance in relation to GFP expression. (A-D) Single planes of confocal z-stack depicting representative phenotype (single vs. multi AC, bottom left of each image) of fluorescence alone (left; AC, magenta, expressing *cdh-3*>mCherry::moeABD and BM, green) and DIC overlay (right). (E) Sina plots of GFP-tagged TF levels, defined as the mean gray value of individual AC nuclei following TF-RNAi knockdown. In this and all other figures, statistical significance as compared to empty vector controls is denoted as an open black circle and here represents a p-value of $< 1\times10^{-6}$ by Student's t test ($n \geq 50$ animals per treatment). Asterisk (*) denotes quantification of animals with ACs (single or multi) that failed to invade.

Figure 5 Medwig-Kinney et al., 2019

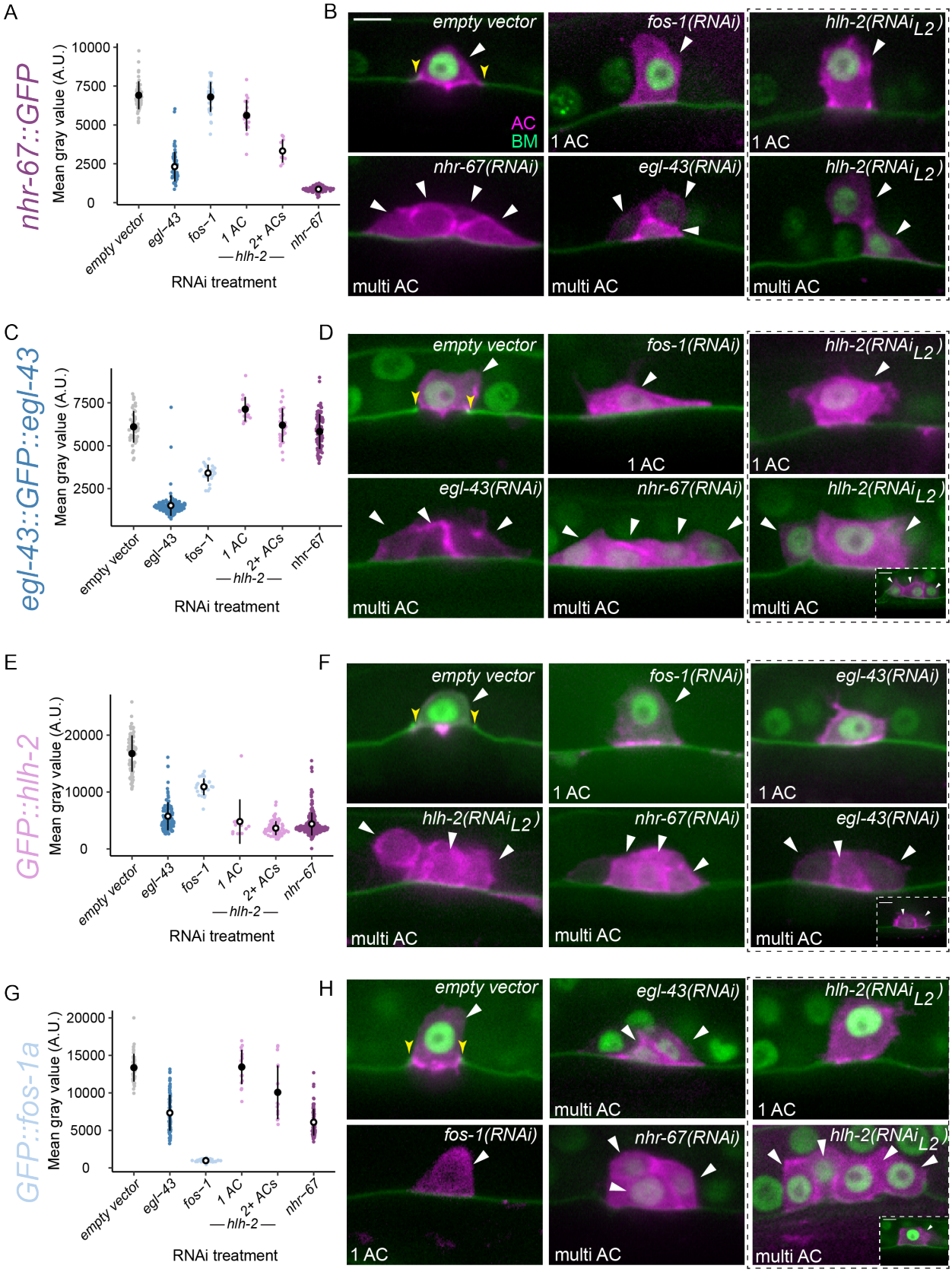


Fig. 5. Regulatory interactions among pro-invasive TFs at endogenous loci. (A, C, E,G) Sina plots of GFP-tagged TF levels, defined as the mean gray value of individual AC nuclei, following RNAi perturbation (n ≥ 25 animals per treatment). (B, D, F, H) Single planes of confocal z-stack depicting representative phenotypes (single vs. multi AC, as noted bottom left of each image) following TF depletion by RNAi, as denoted in the upper right of each image. Dashed boxes indicate representative images of single vs. multi AC phenotypes as a result of specific TF-RNAi treatments.

Figure 6 Medwig-Kinney et al., 2019

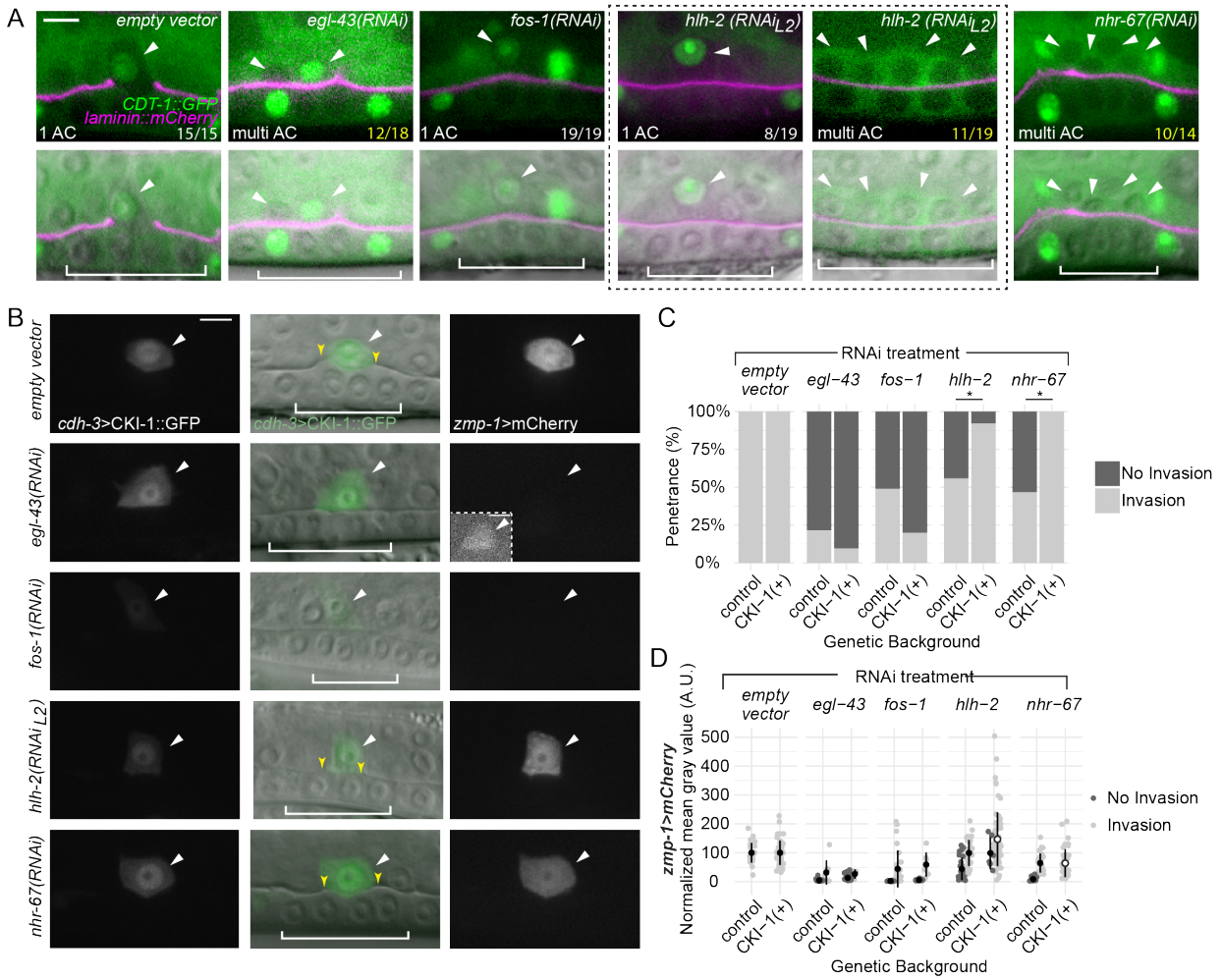


Fig. 6. HLH-2 depletion is rescued by induced G1/G0 arrest while EGL-43 has a cell cycle-independent pro-invasive role. (A) Localization of cell cycle state reporter, *cdt-1*>CDT-1::GFP (green) and BM (*lam-1*>LAM-1::mCherry, magenta) in empty vector control (left) as compared to TF-RNAi depletions (top: fluorescence alone, bottom: DIC overlays). White and yellow fractions (bottom right of each fluorescence micrograph) indicate number of occurrences of nuclear vs non-nuclear *cdt-1*>CDT-1::GFP, respectively. (B) Single confocal plane depicting localization of *cdh*-3>CKI-1::GFP (left), DIC overlay (middle), and *zmp-1*>mCherry expression (right) in empty vector control (top) and TF-RNAi depletions as indicated (left). (C) Stacked bar graph depicting penetrance of AC invasion defects in control animals lacking *cdh-3*>CKI-1::GFP as compared to CKI-1(+) animals. Asterisk (*) denotes statistical significance between control and CKI-1(+) animals and represents a p-value < 0.00001 by Fisher's exact test

(n ≥ 27 animals per treatment). (B) Sina plots depicting quantification of *zmp-1>mCherry* reporter levels in control and CKI-1(+) animals. Statistical significance represents a p-value of < 0.01 by Student's t test (n ≥ 27 animals per treatment).

Figure 7

Medwig-Kinney et al., 2019

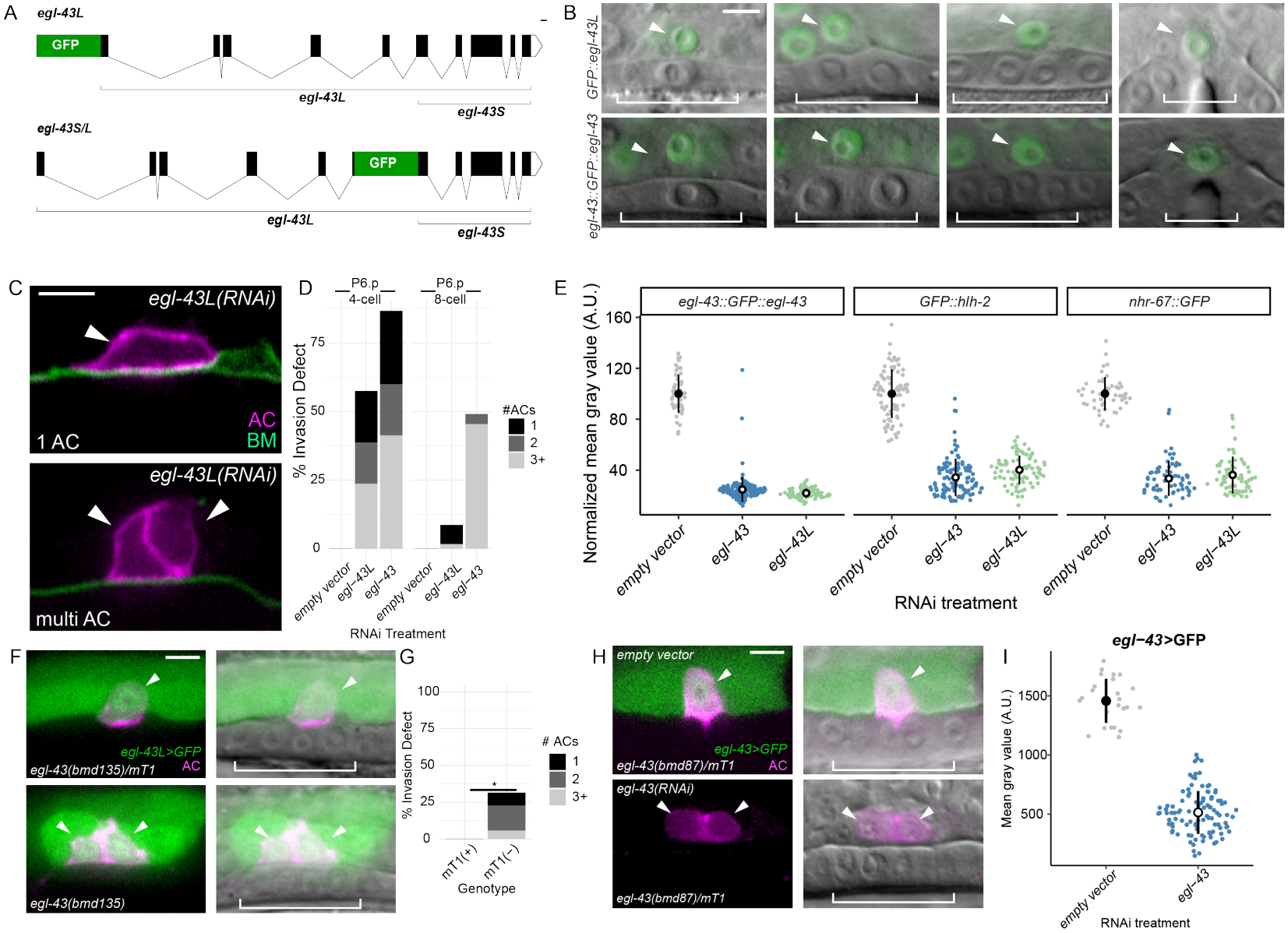


Fig. 7. Both isoforms of *egl-43* function redundantly to regulate AC invasion. (A) Schematics (via http://wormweb.org/exonintron) of GFP insertion into the *egl-43* locus to tag the long (top) or both isoforms (bottom). Scale bar, 100bp. (B) DIC overlaid with single confocal plane for *GFP::egl-43L* (top) and *egl-43::GFP::egl-43* (bottom). (C) Single confocal plane depicting single AC (top) and multi AC phenotypes (bottom) following *egl-43L(RNAi)* treatment. The AC-specific membrane marker (magenta) and BM marker (green) are overlaid in each micrograph. (D) Stacked bar chart showing penetrance of invasion defects, binned by number of ACs. (E) Sina plots of GFP-tagged TF levels, defined as the mean gray value of individual AC nuclei, following RNAi perturbation ($n \geq 25$ animals for each treatment) between RNAi treatments targeting *egl-43L* and *egl-43*. (F) Single plane of confocal z-stacks depicting representative micrographs of control (top) versus *egl-43(bmd135)* animals expressing *egl-43L>GFP* (green) and an AC reporter (magenta). (G) Stacked bar graph depicting penetrance of AC invasion defect. Asterisk (*) denotes statistical significance between control and mT1(-) animals and represents a p-value < 0.0008 by Fisher's exact test ($n \geq 25$ animals per treatment). (H) Single plane depicting fluorescence micrographs (left) and DIC overlays (right) of balanced *egl-43(bmd87)/mT1* animals expressing endogenous *egl-43>GFP* (green) and an AC reporter (magenta). (I) Sina plots of *egl-43>GFP* levels, defined as the mean gray value of individual ACs, following RNAi perturbation ($n \geq 25$ animals per treatment).

Figure 8 Medwig-Kinney et al., 2019

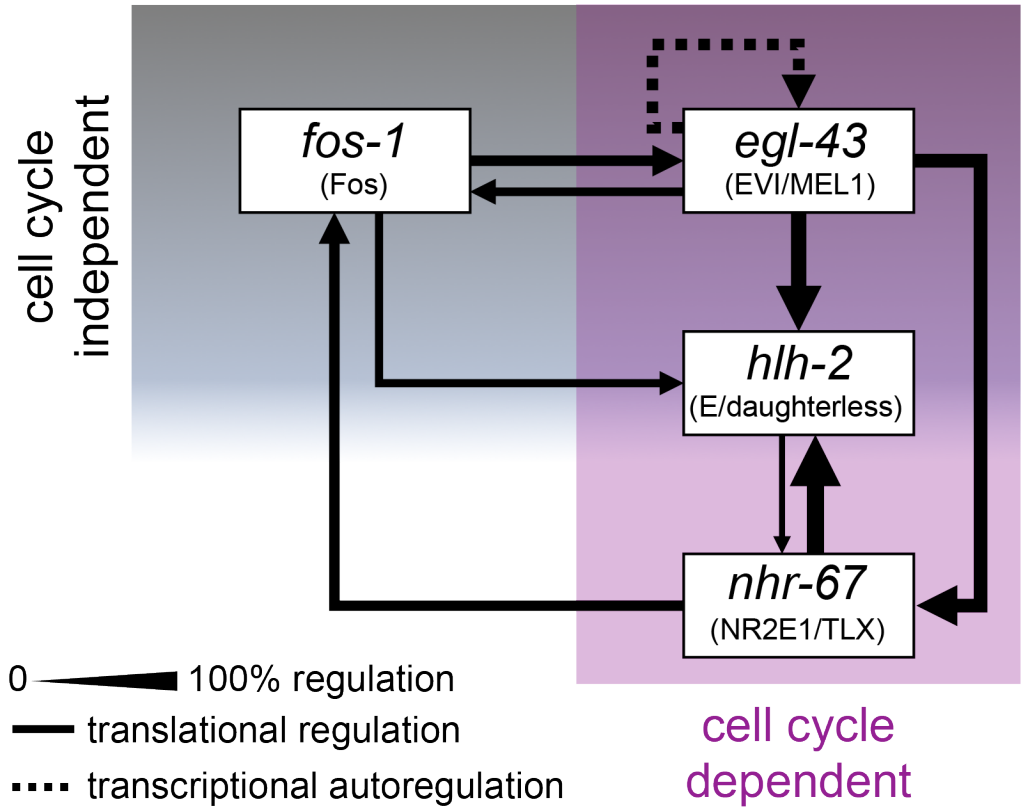


Fig. 8. Summary model of the GRN coordinating AC invasion. Network inference predicts cell cycle independent (gray shaded region) and dependent (purple shaded region) sub-circuits. The thickness of the arrow corresponds to strength of the regulatory interaction, determined by normalizing perturbation data to that of both positive and negative controls. Solid lines refer to regulation of protein levels, dotted line refers to transcriptional regulation.