國立臺灣大學醫學院分子醫學研究所
碩士論文

Institute of Molecular Medicine
College of Medicine
National Taiwan University
Master's Thesis

於秀麗隱桿線蟲中探討調控精子之小RNA的生成機制

Decoding the Mechanisms Governing Small RNA Biogenesis Essential for Sperm Function in *Caenorhabditis elegans*

吳念祖
Nien-Tsu Wu

指導教授：蔡欣祐 博士
Advisor: Hsin-Yue Tsai, Ph.D

民國 113 年 6 月
June 2024

doi:10.6342/NTU202403521

國立臺灣大學碩士學位論文
口試委員會審定書

MASTER's THESIS ACCEPTANCE CERTIFICATE
NATIONAL TAIWAN UNIVERSITY

於秀麗隱桿線蟲中探討調控精子之小RNA的生成機制
Decoding the Mechanisms Governing Small RNA Biogenesis Essential for Sperm Function in *Caenorhabditis elegans*

本論文係吳念祖（R11448002）在國立臺灣大學分子醫學研究所完成之碩士學位論文，於民國 113 年 6 月_20 日承下列考試委員審查通過及口試及格，特此證明。

The undersigned, appointed by the Institute of Molecular Medicine on 20/06/2024_have examined a Master Thesis entitled above presented by Nien-Tsu Wu (R11448002) candidate and hereby certify that it is worthy of acceptance.

口試委員 Oral examination committee:

林劭品 詹世鵬 _____

（指導教授 Advisor）

蔡欣祐 _____ _____

系（所、學位學程）主管 Director: 潘俊良

doi:10.6342/NTU202403521

誌謝

這篇論文的完成離不開許多人的指導、支持和鼓勵，在此我想向他們表達我最衷心的感謝。 首先我要感謝我的指導教授蔡欣祐老師，在我的研究過程中老師提供了寶貴的指導、深刻的反饋和不懈的支持，有老師的專業知識和實質建議，我的實驗才能順利進行。我也非常感謝林劭品教授及詹世鵬教授願意擔任我的論文指導委員，老師們建設性的提問使我更加了解自己的研究內容及方向。

特別感謝學妹 Veronica Kao 幫忙建構報導基因，學弟余佳祐幫忙養線蟲，以及其他實驗室夥伴在我遇到問題時都會給予建議，他們的支持和友誼使這段充滿挑戰的旅程更加愉快和有意義。此外，我也要感謝徐立中教授及張智芬教授提供抗體及試劑使我有更多不同的實驗技術能操作，使我的研究得以順利進行。

最後，我深深感謝我的家人和朋友，感謝家人們金援上以及朋友們精神上的支持，他們無條件的愛、耐心和鼓勵是我學術道路上持續不斷的力量源泉，也謝謝我家的貓陪我讀論文、做報告。

doi:10.6342/NTU202403521

摘要

小 RNA 在各種物種的精子功能中起著至關重要的作用。在秀麗隱桿線蟲（Caenorhabditis elegans）中，缺乏精子相關的小 RNA「alg-3/4 26G 小 RNA」會導致線蟲在 25 攝氏度時不育，然而，大多數先前的研究集中在缺乏 alg-3/4 26G 小 RNA 的病理影響上，其上游調控機制仍不清楚。因此我們的研究深入探討了 26G sRNA 的生物合成機制，針對類 zc3h12a 核糖核酸酶 NYN-3 如何辨識 mRNA 模板並進行切割進行研究。欲探討 26G 小 RNA 生成機制，我們分析了 ALG-3 結合的 mRNA 中的序列元素和轉錄起始位點（TSS），提出 NYN-3 的識別並不依賴特定序列，而是依靠結合轉錄起始位點的蛋白進行辨識。

根據文獻探討，我們發現線蟲中真核翻譯起始因子 IFE-1 作為結合轉錄起始位點的蛋白，非常有可能協助 NYN-3 辨認 mRNA 模板。IFE-1 是人類真核翻譯起始因子 EIF4E 的直系同源物，主要表達於雄性生殖細胞系統中，並表現於雄性生殖腺的近端區域，這與 NYN-3 和 ALG-3 的基因表現相一致。我們的研究結果發現 IFE-1 在雄性生殖腺的近端區域與 NYN-3 短暫相互作用，推斷 IFE-1 能幫助 NYN-3 識別 mRNA 進而產生 26G sRNA 合成所需的 RNA 模板。

關鍵詞：秀麗隱桿線蟲、精子生成、轉譯起始因子、生殖細胞發育、26G 小 RNA、ALG-3/4

doi:10.6342/NTU202403521

Abstract

Small RNAs play a crucial role in sperm function across various species. In *Caenorhabditis elegans*, the absence of sperm-related small RNAs, *alg-3/4* 26G small RNAs, results in infertility at 25 degrees Celsius. Most prior studies focused on the pathological effects of missing *alg-3/4* 26G small RNAs, the upstream regulatory mechanisms have remained unclear. Our research delves into the mechanisms of 26G sRNA biogenesis, particularly the targeting rules of the zc3h12a-like ribonuclease NYN-3, which is essential for the cleavage of mRNA templates required for 26G sRNA synthesis. We investigate the sequence elements and transcription start sites (TSSs) in ALG-3-bound mRNAs, proposing that NYN-3 recognition is independent of specific sequences, but rather involves interactions with TSS-binding proteins. We further identify the eukaryotic translation initiation factor IFE-1 as a potential TSS-binding protein that may assist NYN-3 in 26G sRNA biogenesis. IFE-1, an ortholog of human EIF4E, is predominantly expressed in the male germline and localizes to the proximal region of the male gonad, aligning with the expression patterns of NYN-3 and ALG-3. Our findings demonstrate that IFE-1 transiently interacts with NYN-3 at the proximal region of the male gonad, suggesting a collaborative role in recognizing target mRNAs for 26G sRNA production.

Key words: *C. elegans*, Spermatogenesis, Translation initiation factors, germline development, 26G small RNAs, ALG-3/4

doi:10.6342/NTU202403521

Table of Contents

doi:10.6342/NTU202403521

doi:10.6342/NTU202403521

doi:10.6342/NTU202403521

Chapter 1. Introduction

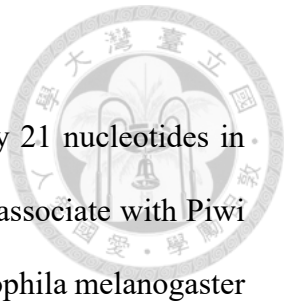
1-1 The role of small RNAs in germline development

Small RNAs (sRNAs), typically 20-30 nucleotides in length, are pivotal regulators of gene expression. Central to the functionality of sRNAs in gene regulation is the Argonaute protein, which shields the sRNAs from the external environment and keeps sRNAs from degradation [1]. When sRNAs form complexed with Argonaute protein, they can recruit regulatory machinery to enhance or suppress gene expression [2-5]. These molecules influence various biological processes, including germline development across diverse species. The regulation of sRNAs in germline development is highly conserved in many species [6].

In *C. elegans*, spermatogenesis is more heavily regulated by sRNAs compared to oogenesis due to the unique challenges and requirements of sperm development. Spermatogenesis involves continuous and rapid cell divisions, extensive chromatin remodeling, and the formation of highly specialized sperm cells [7]. In contrast, oogenesis involves fewer cell divisions and a more prolonged developmental process [8]. The high turnover and specific developmental gene expression in spermatogenesis demand tight regulatory control by sRNAs to ensure fidelity and functionality [9, 10].

In terms of sRNAs involved in germline development in *C. elegans*, the well-studied classes include microRNAs (miRNAs), small interfering RNAs (siRNAs) and Piwi-interacting RNAs (piRNAs). Among them, miRNAs and siRNAs exhibit boarder roles in different tissues at different stages, while piRNAs are particularly noteworthy for their crucial role in sperm development.

doi:10.6342/NTU202403521



1-2 The regulation of piRNAs during spermatogenesis

In *C. elegans*, piRNAs are a class of small non-coding RNAs, typically 21 nucleotides in length with a 5′ uracil (21U-RNAs) length. piRNAs were first found to associate with Piwi proteins to form piRNA-induced silencing complexes (piRISCs) in Drosophila melanogaster [11, 12]. In C. elegans, piRNAs interact with the PIWI homolog, PRG-1 (Piwi-related gene 1), to from piRISCs. piRISCs target and silence transposon elements in the genome, thereby preventing genome instability, especially during the rapid cell divisions in spermatogenesis [13, 14]. piRNAs that silence transposon elements mostly originate from individual loci in two clusters on Chromosome IV, each demarcated by a GTTTC consensus motif [15]. The role of piRNAs in transposon silencing during spermatogenesis has been extensively studied during the past few decades.

A novel role of piRNA activating gene expression during spermiogenesis has been identified recently. Spermiogenesis refers to the final stage of spermatogenesis, a process round spermatid become mature and motile spermatozoa. During spermatogenesis, mRNAs required for spermiogenesis are transcribed and stored in a translationally inactive state until needed for translation [16]. In mouse, piRNAs interact with MIWI and target to mRNAs crucial for spermiogenesis, allowing the recruitment of translation initiation factors to promote the translation of target mRNA [17].

1-3 The *alg-3/4* 26G small RNAs specifically regulate spermiogenesis

Unlike the absence of piRNA in mice and flies, the lack of 21U-RNAs in C. elegans does not result in severe defect in sperm function. However, another class of small RNAs, the *alg-3/4* 26G small RNAs (26G sRNAs), has been shown play a role in the regulation of sperm

doi:10.6342/NTU202403521

development, albeit independent of transposon silencing mechanisms. 26G sRNAs are approximately 26 nucleotides in length and contain 5'-monophosphorylated-guanosine. The antisence 26G sRNAs form complexes with Argonaute proteins, ALG-3 and ALG-4 (ALG-3/4), which are specifically expressed during sperm development. 26G sRNAs complexed with ALG-3/4 target to genes important for spermiogenesis and enhance the recruitment of RNA polymerase II to promote transcription [4, 5]. Besides, 26G sRNAs are reported to facilitate the production of 22G small RNAs, which complex with CSR-1 and WAGO-class Argonautes to mediate mRNA activation and silencing, respectively [3]. All those studies suggest 26G sRNAs can transcriptionally or post-transcriptionally regulate sperm-specific gene expression in both up-regulating and down-regulating fashions.

Moreover, 26G sRNAs are found to promote targeted gene expression. In *alg-3/4* mutants, most of the proteins targeted by 26Gs sRNAs are decreased and few of them are increased at 25℃ [4]. The expression of major sperm proteins (MSP), which is important for sperm motility, are also found to be decreased without the presence of 26G sRNAs [5]. All those finding indicate that 26G sRNAs has greater impact on activating rather than suppressing targeted gene expression.

Given that sperm maturation is relied on 26G sRNAs, the role of 26G sRNAs in spermiogenesis has been emphasized. During spermiogenesis, spermatids develop a pseudopod structure that is important for their motility, while they exhibit abnormal structure at 25℃ when 26G sRNAs is absent [3]. Furthermore, the missing of 26G sRNAs result in defective sperm motility and a temperature-sensitive sterile phenotype, which worms have reduce progeny at 20℃ and no progeny at 25℃ [3]. Unlike piRNAs target to both transposon

doi:10.6342/NTU202403521

and sperm-specific genes, 26G sRNAs exclusively target to sperm-specific genes but transposons, indicating s specialized role of 26G small RNAs in sperm-specific gene expression [3].

1-4 The biogenesis of *alg-3/4* 26G small RNAs

Unlike well-studied piRNAs, the biogenesis of 26G sRNAs remain largely unknown. Previous studies have found 26G sRNAs were synthesized by RNA-dependent RNA polymerase (RdRP) RRF-3, which uses cleaved mRNA as template to synthesis complementary RNA strands [18]. The complementary RNA strand form double-stranded RNA (dsRNA) with the original templated, and the dsRNA is subsequently processed by DCR-1 to generate 26G sRNA [19]. However, detail mechanism how RRF-3 target to the RNA template and recruit other proteins to generate 26G small RNAs is still under investigation.

Recent studies in our lab have found a zc3h12a-like ribonuclease, NYN-3, participate in the cleavage of the mRNA targeted by 26G sRNAs (*alg-3/4* targeted mRNA). In NYN-3 catalytic defective mutant, the absence of 26G sRNAs, along with NYN-3's recognition of the targeted mRNA prior to RRF-3, indicates an indispensable role for NYN-3 in the biogenesis of 26G sRNAs [5]. NYN-3 recognize 16-30% of 5' end of *alg-3/4* targeted mRNAs and cleave the mRNAs to generate template RNAs for 26G sRNA synthesis [5]. Although it has been shown RRF-3 targets to *alg-3/4* targeted mRNAs after NYN-3 cleavage, whether NYN-3 cleavage is sufficient for RRF-3 to recognize the template RNA remains mysterious.

doi:10.6342/NTU202403521

1-5 The potential mechanism NYN-3 targeting *alg-3/4* 26G small RNAs

In mouse, the ZC3h12a protein, Regnase-1, recognize double stem-loop and AU-rich elements on 3’ untranslated region (UTR) of mRNA and cleave the translationally active mRNA only when interacting with UPF-1. Interestingly, we found stem-loop structure and AG-rich motifs on 5’ end of *alg-3/4* targeted mRNAs. However, less than 10% of the *alg-3/4* targeted mRNAs exhibit AG-rich motifs, and most of RNA sequence include stem-loop structures. Therefore, stem-loop structure and AG-rich motifs are unlikely to be sufficient for NYN-3 recognition, suggesting that other elements on *alg-3/4* targeted mRNAs or associated proteins might be involved.

To investigate the rules of NYN-3 targeting *alg-3/4* targeted mRNAs, we correlated the potential position of transcription start sites (TSSs) on *alg-3/4* targeted mRNAs from our previous study and 26G sRNAs. Surprisingly, we discovered a consistent distance between TSSs on *alg-3/4* targeted mRNAs and the 5’ end of 26G sRNAs. Consequently, we hypothesized that the rules governing NYN-3's recognition of alg-3/4 targeted mRNAs likely involve both sequence elements on the targeted mRNA and TSS-binding proteins.

1-6 Eukaryotic translation initiation factor in *C. elegans*

To study the potential TSS-binding proteins involved in 26G sRNAs biogenesis, we investigated cap-binding proteins involved in the regulation of spermatogenesis in *C. elegans*. Extensive literature research led us to focus on IFE-1, an ortholog of human EIF4E (Eukaryotic translation Initiator Factor 4E). Human EIF4E binds to the 5' cap structure of mRNAs, facilitating the recruitment of other translation initiation factors and ribosomes,

doi:10.6342/NTU202403521

thereby promoting the initiation of protein synthesis. This process is essential for the efficient translation of most eukaryotic mRNAs [20].

There are five orthologs of human EIF4E in C. elegans. Among the five EIF4E homologs in C. elegans, only IFE-1 and IFE-3 are germline-expressed. [21] [22]. Depletion *ife-1* causes fertilization defects in sperm, while *ife-3* depletion results in defective meiosis and growth in oocyte [21, 22]. IFE-1 interacts with PGL-1, a component of germ granule, to mediate translational activation while IFE-3 bind translation repressor IFET-1 to mediate translational repression [23]. Those findings emphasis on a specified role of IFE-1 in activating translation during spermatogenesis.

1-7 The potential roles of IFE-1 in 26G small RNA regulation

In previous study, *ife-1(bn127)* null mutants have lower progenies at 20℃ and almost no progeny at 25℃, displaying a phenotype similarly to *alg-3/4* mutants and NYN-3 null mutants, indicating a potential role of IFE-1 in 26G sRNAs biogenesis [3, 5, 24]. Moreover, IFE-1 localized primarily to the proximal region in the male gonad, consistent with the expression patterns of both NYN-3 and ALG-3 in the male germline [3, 5, 22]. However, IFE-1 and ALG-3 are mainly expressed in P granule, the germ granule where small RNAs and RNA-binding proteins regulate gene expression, while NYN-3 is expressed in the cytoplasmic region of germline. Given that IFE-1 is found to activate translation during spermatogenesis and IFE-1 and ALG-3 are both expressed in the P-granule during the same stage of spermatogenesis, IFE-1is likely to involve in the ALG-3/4 26G small RNA pathway to regulate sperm-related gene expression.

doi:10.6342/NTU202403521

Chapter 2. Materials and Methods

2-1 Strains

The Bristol N2 strain was used as the wild type for hermaphrodites , while the male of the Bristol N2 strain containing *fog-2(q71)* mutation served as the wild type for males in all experiments, unless specified otherwise. *msd-1::gfp* reporter was designed by Hsin Yue Tsai and constructed by Veronica Kao. Strains related to NYN-3 were described in our previous studies [5]. *ife-1::2x flag* strain was established by tagging 3' UTR with 2x FLAG tag using CRISPR/Cas9 system. *ife-1::2x flag* strain was crossed with *ha::nyn-3 or ha::nyn-3 (ht044)* to study the interaction between IFE-1 and NYN-3. *ife-1(tm4249)* strain was obtained from the National BioResource Project (NBRP) in Japan. All strains related to male studies were cross with *fog-2(q71)* strain to obtain 50% population of males and they were maintained on normal growth medium (NGM) plates at 20°C with E. coli strain OP50 as previous described [25].

The strains used in this study are listed as follows: Bristol N2, CB4108 ([*fog-2(q71)/V*]), HYT054 ([*ha-c29f5.3(CDD)/II; fog-2(q71)/V*)], HYT193 ([*ha::nyn-3(ht044); rol-6/II; ife-1:2xflag/III; fog-2(q71) V*]), HYT194 ([*ife-1(tm4249)/II*]), HYT213 ([*ha-nyn-3(c29f5.3)/II; 2x flag::alg-3/III; fog-2(q71)/V*]), HYT214 ([*msd-1:gfp/V; fog-2/V*]),

2-2 Small RNA library preparation

To prepare small RNA library, 2 x 10^5 male worms of *msd-1::gfp; fog-2(q71)* or *fog-2(q71)* were collected with M9 buffer and resuspended in NucleoZol. Following the NucleoZol RNA isolation protocol (Macherey-Nagel), total RNA and small RNA-enriched fraction were subsequentially extracted. For further purification, 10 ug of small RNA-enriched faction was

doi:10.6342/NTU202403521

loaded onto 15% (w/v) denaturing polyacrylamide gel [1X TBE, 7 M urea, 15% acrylamide (19:1)]. After electrophoresis, small RNAs ranging in size between 20 to 40 nucleotides were visualized by staining with SYBR Green II RNA gel stain (Thermo Fisher Scientific). The desired RNA bands were then excised from the gel. The excised RNA was subject to overnight diffusion elution in Elution buffer (10 mM Tris-HCl PH 7.5, 300 mM NaCl, 1mM EDTA). The eluted RNA was recovered by isopropoanol precipitation, involving the addition of an equal volume of isopropanol and storage at -20 °C overnight.

For cloning, selected small RNAs were ligated with 20 pmole of 3' linker by 20 units of T4 RNA ligase (Takara) in 1x ligation buffer (0.5M Tris-HCl PH7.5, 0.1 M $MgCl_2$, 0.1M DTT) with 0.1 % BSA, 10% DMSO, 20 units of SUPERase-In RNAase inhibitor (Invitrogen) added for 2 hours at 25°C. Ligated small RNAs were separated from non-target RNA fractions by running 7M urea /15% (w/v) polyacrylamide gel in TBE buffer, followed by gel-extraction using same methods as described. Subsequently, 3' ligated small RNAs were ligated with 20 pmole of 5' linker by 13 units of T4 RNA ligase (New England Biolabs) in 1x T4 ligase reaction buffer (50 mM Tris-HCl pH 7.5, 10 mM MgCl2, 1 mM ATP, 10 mM DTT) with 20% PEG8000, 0.001M ATP, 10% DMSO and 8 units of Murine RNase inhibitor (NEB) added. The sequences of 3' and 5' linkers are listed in Table I. Subsequently, 5' ligated RNAs were separated from non-target RNA fraction by 20 µl MyONE Silane beads (Invitrogen) following eCLIP protocol described in 2016 [26].

Ligated small RNAs were subjected to reverse transcription with designed primer following SuperScript^{TM}III Reverse Transcriptase Manual (Invitrogen) (Table I). To prepare small RNA library, cDNA samples were subjected to two rounds of polymerase chain reaction

doi:10.6342/NTU202403521

(PCR). First, cDNA samples were amplified by Q5 polymerase (NEB) with 5' adaptors P7 and 3' adaptor P5 as primers, in final volume of 20 µl. The first round of PCR cycle was as follows: 98°C (30s) for initial denature. 6 cycles of 98°C (15s), 68°C (30s), 72°C (40s) and 4 cycles of 98°C (15s), 72°C (60s), After the first round PCR, PCR products were diluted in 1:10 to determine the cycle number of second round PCR. The cycle number of second round PCR was determined from Ct value in quantitative PCR (qPCR) and the amplicons were checked using 8% (w/v) polyacrylamide gel with PAGE to ensure no bulged products were generated after PCR. For the second round of PCR, PCR products were also amplified by Q5 polymerase (NEB) with 5' adaptor P7 and 3' adaptor P5 as primers, but in final volume of 100 µl. The second round of PCR cycle was as follows: 98°C (30s) for initial denature. 6 cycles of 98°C (15s), 68°C (30s), 72°C (40s) and ((Ct value-3) - 6) cycles of 98°C (15s), 72°C (60s). For two rounds of PCR, the primers The PCR products were concentrated by ethanol precipitation and loaded into 6% (w/v) polyacrylamide gel. After electrophoresis, the final amplicons (150 to 160 bp) were excised, eluted, and used as the input to check the presence of target small RNAs.

2-3 Verification of *gfp* small RNA in reporter strain

Target small RNAs were detected by PCR using designed primers, one corresponding to the target gene and another to the adaptor P7. The primers used to detect different target small RNAs were listed in Table II. For PCR, cDNA from *fog-2* or *msd-1::gfp* library were mixed with 5 µM forward primer, 5 µM reverse primer and 2X taq mastermix (Q-Amp™). The PCR products were analyzed in 8% (w/v) polyacrylamide gel with PAGE to verify the size of amplicons. The amplicons in polyacrylamide gel were visualized by staining with SYBR

doi:10.6342/NTU202403521

Green II RNA gel stain (Thermo Fisher Scientific). *mir-66* and *f59a6.2* were used as positive control for both wild type and reporter strain.

The target amplicons of *gfp* small RNA were diffused in Elution buffer and precipitated by isopropanol precipitation. Subsequently, the target amplicons were conjugated with cloning vector using CloneJet PCR Cloning Kit (Thermo), followed by sanger sequencing. The primer used for sanger sequencing was based on the 5' end of cloning vector, as illustrated in Figure 1E and its sequence listed in Table II.

2-4 Detection of alternative transcripts in males and hermaphrodites

To confirm the presence of alternative transcript is biased by different sex, total RNAs were extracted from 1000 hand-picked male worms under *fog-2(q71)* background, or extracted from N2 hermaphrodites, following the user manual of NucleoZol RNA isolation (Macherey-Nagel). 5 μg of total RNAs and 0.5 μg of oligo(dT) were applied in final volume of 20 μl for reverse-transcription (RT), following MMLV Reverse Transcription Manual (Protech).

To identify alternative transcripts, forward primers for *f59a6.2* were designed based on the newly identified transcription start site (TSS) or TSS from previous studies (Table III) [5, 27]. For polymerase chain reaction (PCR), 0.25 μl of RT product, 5 μM of forward and reverse primers and 2X taq mastermix (Q-Amp^{TM}) were mixed for DNA amplification. PCR products were amplified by 35 or 40 cycle to ensure the presence or absence of target expression.

2-5 Brood size test

Worms of N2 or *ife-1(tm4249)* were maintained on NGM plates with OP50 at either 20°C or 25°C to confirm whether *ife-1* null function results in temperature-sensitive sterile phenotype.

doi:10.6342/NTU202403521

Worms were singled out at L4 stage and one worm was plated on one NGM plate for ten repetitions to examine the brood size. To observe daily changes in the brood size, individual worm was transferred to a new plate until no worm progeny was observed. The whole counting process took about five to six days. Statistical analysis was conducted by Prism6 and p-values were calculated using the ordinary one-way ANOVA followed by an unpaired t-test.

2-6 Worm lysate preparation

All worms subjected to lysate extraction were under *fog-2(q71)* backgrounds. Synchronized males were enriched by filtering through a 35-μm mesh (Spectrum Labs, New Brunwick, USA), collected using M9 buffer and resuspended in an equal volume of immunoprecipitation (IP) buffer (25mM HEPES-KOH PH 7.5, 10 mM KOAc, 110 mM KCl, 2 mM $MgCl_2$, 0.1% NP-40, 1 mM Dithiothreitol) with Protease inhibitor Cocktail (1x cOmplete^{TM} ULTRA Tablets, Mini, EASYpack, Roche) and SUPERaseIn RNAase inhibitor (1 μl per 2 ml lysates, Invitrogen). Worm lysates were homogenized by vigorously shacking with glass beads (Sigma, 425-600 μm) using homogenizer (Clubio) with conditions of 3,500 rpm, 10 seconds, 4 cycle and 1 minute interval on ice. The cuticles were removed by two rounds of centrifugation at 10,000 rpm for 10 minutes each. The supernatant containing target proteins can further be used for western blot or immunoprecipitation assay.

2-7 Immunoprecipitation

The protein content of worm lysates was quantified using the Bradford assay. 1 mg of total lysates were incubated with mouse 1 μl mouse anti-FLAG (Sigma Aldrich, F1804) or 1 μl rabbit anti-HA (Cell signaling Technology,3724S) for 1 hour, followed by adding 20 μl

doi:10.6342/NTU202403521

Protein A/G PLUS-Agarose (Santa Cruz) and incubation for another 1 hour. During the incubation, protein lysates and beads were tumbled end over end at 4 degrees Celsius in a tube rotator. After the incubation, agarose beads conjugated with target proteins were centrifuged at 5,000 rpm for 1 minute to separate immunoprecipitated proteins with flow-through fractions. Subsequently, beads were washed with ice-cold 1 ml IP buffer for three times.

2-8 Western blot

Protein samples from worm lysate were prepared by mixing with SDS loading dye (0.4% SDS, 0.25M Tris-HCl PH6.8, 30% glycerol, 2 mg/ml bromophenol blue, 0.16% mercapotoethanol) and 10 mM Dithiothreitol (DTT), followed by heating at 95 degrees Celsius for 10 minutes. IFE-1::2X FLAG and HA::NYN-3 proteins were separated at expected size of ~26 kDa and ~45 kDa each by 15% sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Subsequently, target proteins were transferred from polyacrylamide gel to a PVDF membrane (Merck) with Trans -Blot semi-dry transfer cell at constant voltage of 20V at room temperature for 1 hour. The transferred membrane containing target proteins were blocked with 5% non-fat milk in PBST for 1 hour. Membrane containing 1::2X FLAG or HA::NYN-3 were separated based on their respective protein size, followed by incubation with mouse anti-FLAG (1:1000) (Sigma Aldrich, F1804) or rabbit anti-HA (1:1000) (Cell signaling Technology,3724S) in PBST overnight at 4 degrees Celsius. Next day, membranes were incubated with HRP-conjugated anti-mouse (1:5000, Jackson ImmunoResearch, 115-035-003) or HRP-conjugated anti-rabbit (1:5000, Jackson ImmunoResearch, 113-035-003), respectively. The target proteins were visualized using a

doi:10.6342/NTU202403521

ChemiDoc-It Imager (Analytik Jena company) with Western Lighting Plus ECL (PerkinElmer). Between blocking, antibody incubation and visualization, membranes were washed with PBST for 15 minutes in three repeats.

2-9 Immunofluorescence staining

Adult male gonads of *ife-1::2x flag; ha::nyn-3(c29f5)* and *ife-1::2x flag; ha::nyn-3(ht044)* were released in EBTA buffer (25 mM HEPES-NaOH, PH 7.4, 118 mM NaCl, 48 mM KCl, 2 mM EDTA, 0.5 mM EGTA, 0.1% Tween-20 and 20 mM sodium azide) by decapitating worms on a slide. Extruded gonads were fixed with 3.7% formaldehyde for 5 minutes. Subsequentially, samples are frozen by placing on the aluminum block in -80℃ refrigerator for >5 minutes and permeabilized by freeze-cracking and 1-minute incubation with 100% ice-cold methanol. Permeabilized gonads were blocked with 5% bovine serum albumin (BSA) (Sigma Aldrich, B4287) in PBST (1x PBS, 0.1% Tween 20) for 30 minutes. After blocking, slides were washed 10 minutes for three times by moving slides to fresh Coplin jars of PBST. For staining samples were incubated with mouse anti-FLAG (1:50) (Sigma Aldrich, F1804) and rabbit anti-HA (1:50) (Cell signaling Technology,3724S) in PBST overnight at 4 degrees Celsius. Next day, slides were washed 10 minutes for three times, followed by staining with Alexa Fluor 555-conjugated goat anti-mouse (1:1000) and Alexa Fluor 488-conjugated donkey anti-rabbit IgG (1:1000) in PBST overnight at 4 degrees Celsius. Fluoroshield mounting medium with DAPI (Abcam, ab104139) was used for mounting and staining the nucleus in male gonads. More details are described in protocols from previous study [28]. Confocal images were acquired using a Zeiss LSM700 confocal microscope. The green

doi:10.6342/NTU202403521

fluorescence (GFP), red fluorescence (RFP) and transmitted light channels were used during image acquisition. The images were processed using Axiovision 4.8 and Zen 2010.

2-10 Proximity ligation assay (PLA)

Adult male gonads of *ife-1::2x flag; ha::nyn-3* and *ife-1::2x flag; ha::nyn-3(ht044)* were prepared using the same methods in immunofluorescence staining experiment. Reagents used in PLA were provided in Duolink® In Situ Red Starter Kit Mouse/Rabbit (Sigma). Permeabilized gonads were blocked with Duolink Blocking Solution for 1 hour at 37°C. Subsequently, gonads were incubated with mouse anti-FLAG (1:50) (Sigma Aldrich, F1804) and rabbit anti-HA (1:50) (Cell signaling Technology,3724S) in PBST overnight at 4°C. Next day, slides were washed 10 minutes for three times by moving slides to fresh Coplin jars of PBST, followed by samples incubated with PLUS and MINUS PLA probes in the Doulink Antibody Diluent (1:5) for 1 hour at 37°C. Slides were washed for 10 minutes three times in triplicate by moving slides to fresh Coplin jars of 1x Wash buffer A (10 mM Tris PH 7.4, 150 mM NaCl, 0.05% Tween 20) at room temperature. Subsequently, samples were incubated with ligase in 1x Duolink Ligation buffer for 30 minutes at 37°C to ligate theprobes. After ligation, slides were washed with 1x Wash buffer A using same methods. Samples were then incubated with DNA polymerase in 1x Duolink Amplification buffer for 100 minutes at 37°C, allowing circular polymerase chain reaction (PCR) amplifies PLA signals. Finally, slides were washed with 1x Wash buffer B (200 mM Tris PH 7.4, 100 mM NaCl). Doulink PLA Mounting medium with DAPI was used for mounting and staining the nucleus in male gonads. Confocal imaging and images processing were described as in immunofluorescence staining experiment. PLA signals in the proximal gonad of male were counted and statistical

doi:10.6342/NTU202403521

analysis was performed with Prism6. P-values were calculated using the ordinary one-way ANOVA with Turkey's multiple comparisons t-test.

doi:10.6342/NTU202403521

Chapter 3. Results

3-1 NYN-3 recognition is sequence-independent and doesn't solely rely on the stem loop structure

To unravel the mechanisms governing 26G small RNA biogenesis, our aim was to develop a reporter system capable of elucidating how these small RNAs are produced and how they modulate gene expression. Previous findings of our studies indicated a down-regulation of major sperm proteins (MSPs) in the absence of 26G small RNA[5]. However, selecting a single representative MSP gene for reporting proved challenging due to the abundance of highly similar MSPs in C. elegans. Consequently, we identified MSD-1 (Major Sperm protein Domain-containing), which co-localizes with MSPs and regulates their polymerization and depolymerization within spermatozoa pseudopods [29]. Proteomic data demonstrated a similar fold decrease in MSP and MSD-1 levels in the absence of 26G small RNAs [4]. Furthermore, our ALG-3-bound 26G small RNA sequencing data indicated that 26G small RNAs targeting MSD-1 are most abundant compared to those targeting other proteins whose levels are up-regulated by 26G small RNAs. Therefore, *msd-1* was selected as the 26G small RNA reporter for further studies.

Our analysis of ALG-3-bound 26G small RNA sequencing and NYN-3 ePAR-CLIP sequencing data revealed that the NYN-3 binding site lies downstream of the two targeted sites for 26G small RNAs on MSD-1 mRNA (Figure 1A). Despite NYN-3 featuring a ZC3h12a-like ribonuclease domain known for cleaving loop regions of stem-loop structures, whether NYN-3 exhibits a preference for recognizing and cleaving such structures remained uncertain.

doi:10.6342/NTU202403521

Remarkably, we identified a high-probability double stem-loop structure near the NYN-3 binding site (Figure 1A). To investigate the influence of secondary structure on NYN-3 cleavage, we engineered a strain with a single-copy insertion of the *alg-3/4* targeted gene, msd-1, fused with a GFP sequence at a position corresponding to both the 26G small RNA template and NYN-3 binding site (Figure 1B). This modification altered both the original secondary structure and sequences, resulting in a completely different sequence and decreased stem-loop potential at the NYN-3 binding site.

To assess the importance of secondary structure for NYN-3 cleavage, we examined the presence of GFP small RNAs in our reporter strain. The detection of GFP small RNAs would indicate that NYN-3 does not rely solely on gene body recognition, and secondary structure is not essential for NYN-3 cleavage.

To investigate the presence of GFP small RNAs, Small RNA libraries were generated for both wild-type and reporter strains, with initial extraction from male young adult worms, size selection to distinguish small RNAs from ribosomal RNAs, ligation with 3' and 5' adaptors, and reverse transcription to create cDNA libraries. Target small RNAs were detected using designed primers, one corresponding to the target gene and one to the adaptor (Figure 1C).

Intriguingly, we observed both wild type and reporter strains containing Mir-66 (a 23-nucleotide-microRNA, which will also be cloneded in the NGS library) and F59A6.2 (the most abundant 26G small RNAs) small RNAs, while only the reporter strain harbored GFP small RNAs (Figure 1D). Sanger sequencing of the expected size of GFP small RNAs confirmed their full match to the original GFP sequence (Figure 1E). These findings suggested that altering the gene sequence and reducing stem-loop potential do not affect

doi:10.6342/NTU202403521

NYN-3 recognition, indicating that NYN-3 cleavage is sequence-independent and not solely reliant on secondary structure.

3-2 The biogenesis of ALG-3/4 26G sRNA likely depends on transcription start sites

If NYN-3 doesn't solely rely on the secondary structure to recognize its target, there might be other factors helping NYN-3 recognition. To find out more clues, we analyzed our previous ALG-3 bound sequencing data to see whether there are more consistent rules for 26G small RNA biogenesis. In our previous studies, we pull down ALG-3 and analyzed the distance between ALG-3-bound small RNAs and *alg-3/4*-targeted mRNAs (Figure 2A). Based on the *alg-3/4*-targeted mRNA, we have identified the potential transcription start sites of the mRNA transcripts for *alg-3/4* 26G small RNA synthesis.

Based on the identified potential transcription start sites (TSSs) for *alg-3/4* 26G small RNA synthesis, we observed that the 5' ends of 26G small RNAs were located approximately 40 nucleotides from the newly identified TSSs for all eight candidate genes (Figure 2B bottom). However, comparison with previously identified TSSs revealed a random distribution of 26G small RNAs relative to these TSSs (Figure 2B top). We hypothesized that this discrepancy may be due to differences in the worm strains used in our study, focusing on male worms, and previous studies, typically employing hermaphrodites.

To investigate potential sex-specific biases in *alg-3/4* target TSSs, we examined the *f59a6.2 transcript,* where the newly identified TSS in males is 25 nt downstream of the previously identified TSS. Two forward primers, along with a common reverse primer, were designed to determine whether the amplicons upstream or downstream of the newly identified *f59a6.2* TSS could be amplified in either male or hermaphrodite cDNAs (Figure 2C). Our results

doi:10.6342/NTU202403521

indicated that both upstream and downstream amplicons of the newly identified TSS were amplified in cDNAs derived from hermaphrodite transcripts, suggesting that the TSS of hermaphrodite *f59a6.2* transcripts is likely the previously identified TSS. In contrast, only the downstream amplicon of the newly identified *f59a6.2* TSS was amplified in male cDNAs, indicating that the TSS of male *f59a6.2* transcripts is not the previously identified TSS and is likely the newly identified TSS from our sequencing analysis.

3-3 IFE-1, a germline-specific homolog of EIF4E, potentially aids NYN-3 in recognizing target mRNAs

Our previous results suggest that NYN-3 is the enzyme that recognizes alg-3/4 targeted transcripts prior to the production of alg-3/4 26G small RNAs. However, neither the primary nor secondary sequence of the NYN-3 binding site is crucial for NYN-3 recognition. Additionally, a constant distance between the 5' end of 26G small RNAs and the male TSS of alg-3/4 targeted mRNAs suggests that protein binding to the TSS also participates in alg-3/4 26G small RNA biosynthesis. We thus hypothesize that NYN-3 binding sites on alg-3/4 targeted transcripts are guided by a TSS-binding protein, which is commonly a cap-binding protein. Strikingly, *ife-1* null mutants exhibited a phenotype similarly to NYN-3 null mutants, including temperature-sensitive infertility corresponding to the sterility observed in *alg-3/4* 26G small RNA mutants. To ensure the phenotypes described in literature is indeed temperature-sensitive sterile, we obtained another *ife-1* allele, *ife-1(tm4249),* from the National BioResource Project (NBRP). The 343 bp deletion (tm4249) spans from nucleotides 217 in exon 2 to nucleotide 559 in exon 3 (Figure 3A). The deletion removes over 50% of the coding region for IFE-1 (115/229 codons), resulting a null mutation. As anticipated, IFE-

doi:10.6342/NTU202403521

1(tm4249) resulted in complete sterility at 25 °C and significantly reduced brood size at 20°C (Figure 3B).

As indicated by previous studies and our protein prediction, human EIF4E possesses a cap-binding pocket comprised of two conserved tryptophan residues and a 4E-binding site on the long tail of the N terminus [30] (Figure 3C). To know whether IFE-1 also contains cap binding pocket, we performed structural prediction using AlphaFold 2 and compare the structural similarity between human EIF4E and IFE-1. Similar to human EIF4E, IFE-1 contains cap binding pocket but differs in symmetrical tails on both C and N terminus (Figure 3D). In IFE-1 null mutant, *ife-1(tm4249)* deletion results in the depletion of cap binding pocket (Figure 3E). Structural predictions reveal that IFE-1 likely harbors the ability to recognize the cap binding capability. Overall, these findings suggests that IFE-1 is likely to be the potential protein interactor of NYN-3.

3-4 IFE-1 transiently interacts with NYN-3 at the proximal region in male gonad

To validate the potential interaction between IFE-1 and NYN-3, we generated functional NYN-3 [*ife-1::2x flag; ha::nyn-3(c29f5)*] and NYN-3 catalytic inactive mutant [*ife-1::2x flag; ha::nyn-3(ht044)* strains under *fog-2(q71)*] strains using CRISPR and genetic crossing strategies. Co-immunoprecipitation experiments in revealed a weak interaction between IFE-1 and NYN-3 in *ife-1::2x flag; ha::nyn-3(c29f5)* strain (Figure 4A). By pulling down IFE-1, a strong band for IFE-1::2x FLAG and a faint band HA:: NYN-3 were observed at the expected sizes of ~26 , ~45 kDa, respectively. In contrast, a strong band HA:: NYN-3 and a faint band IFE-1::2x FLAG were observed when pulling down NYN-3. These findings came

doi:10.6342/NTU202403521

up with a hypothesis of transient interaction between IFE-1 and NYN-3, likely due to a possibility that NYN-3 dissociates from the mRNA template following cleavage. Following this hypothesis, we use NYN-3 catalytic mutant strain, which cannot cleave the RNA template, to further compare with functional NYN-3 for the investigation of IFE-1 and NYN-3 interaction.

To explore the cellular localization and spatial interaction of IFE-1::2X FLAG and HA::NYN-3 in the male germline, we conducted immunofluorescence staining on male gonads. We examined male gonads from both functional NYN-3 [*ha::nyn-3(c29f5); ife-1::2x flag; fog-2(q71)*] and NYN-3 catalytic inactive mutant [*ha::nyn-3 (ht044); ife-1::2x flag; fog-2(q71)*] adult worms grown at 25°C. Consistent with previous findings, both IFE-1::2X FLAG and HA::NYN-3 were predominantly expressed in the proximal region of the male germline (Figure 4B, 4D). Interestingly, HA::NYN-3 was not only distributed in the cytoplasmic area of the germline, as seen in our prior studies, but also formed some P-granule-like puncta around the nucleus (Figure 4C,4E) The signals of IFE-1::2X FLAG and HA::NYN-3 were closely proximal but not fully overlapping in both HA::NYN-3 and catalytic inactive mutant HA::NYN-3 (ht044) strains (Figure 4C, 4E). Given the highly similar expression patterns of IFE-1::2X FLAG and HA::NYN-3 in the proximal region of the male germline, it is plausible that there is a spatial interaction between IFE-1 and NYN-3 in the vicinity of the P-granule.

To visualize the spatial interaction between IFE-1::2X FLAG and HA::NYN-3 in the male germline, we conducted a proximity ligation assay (PLA) on male gonads from the same two

doi:10.6342/NTU202403521

strains. In PLA, when two proteins are in close-proximity, their interaction can be detected using antibodies, and the signal can be amplified. Under the same conditions as the immunofluorescence staining experiment, no PLA signals were observed in wild type male gonad (Figure 4F). Excitingly, we observed PLA signals predominantly in the proximal regions of the male germline in both the functional HA::NYN-3 and catalytic inactive mutant HA::NYN-3 (ht044) strains, indicating the interaction between IFE-1::2X FLAG and HA::NYN-3 occurred in the proximal germline (Figure 4G, 4H). Notably, most PLA signals were located around the nucleus, supporting the hypothesis that IFE-1::2X FLAG and HA::NYN-3 might interact in close proximity to the P-granule (Figure 4G, 4H). Additionally, quantification data revealed significantly more signals in the catalytic inactive mutant HA::NYN-3 (ht044) strain compared to the functional HA::NYN-3 strain (Figure 4J). These results support the hypothesis from the Co-IP experiment, the interaction between IFE-1 and NYN-3 is transient (Figure 4A, 4C).

doi:10.6342/NTU202403521

Chapter 4. Discussion

3-1 Transcription factors and 26G sRNAs both regulate the expression of sperm-specific genes during spermatogenesis

Previous studies have shown that transcription factors (TFs) are specifically expressed at different developmental stage to drive the expression of certain genes that are required for development. TFs can epigenetically reprogram clustered CTCF binding sites into transcriptional start sites, promoting transcription from alternative promoters during spermatogenesis in mammals [31]. Although there are no CTCF sites in *C. elegans*, certain transcription factors are expressed to drive the expression of sperm-specific transcripts during spermatogenesis [32, 33]. In this study, we indeed observed sperm-specific transcription start sites (TSSs) in males undergoing active spermatogenesis. It is likely that the expression of sperm-specific genes, or alg-3/4 targeted mRNAs, is regulated by transcription factors during spermatogenesis.

To gain a comprehensive understanding of the biogenesis of 26G sRNAs it is important to know how *alg-3/4* targeted mRNAs are regulated. To investigate the role of transcription factors in the ALG-3/4 26G small RNA pathway, we can select TFs specifically expressed and involved in spermatogenesis through literature research. Additionally, we can conduct transcription factor enrichment analysis (TFEA) on *alg-3/4* targeted mRNAs to examine which TFs participate in the biogenesis of 26G small RNAs.

doi:10.6342/NTU202403521

3-2 The targeting rules of NYN-3 recognizing *alg-3/4* targeted mRNA remains to be discussed

In this study, we discovered that NYN-3 recognition is sequence-independent and reducing stem-loop potential doesn't affect the production of 26G sRNAs. However, our investigation into NYN-3 recognition focused only on several marker genes, and the secondary structure was not entirely eliminated in the msd-1::gfp reporter. Therefore, we cannot conclusively assert that NYN-3 does not require a stem-loop structure to recognize *alg-3/4* targeted mRNAs. To study whether stem-loop structure is indispensable for NYN-3 recognition of *alg-3/4* targeted small RNAs, we can alter the sequence of *msd-1:gfp* without changing the amino acids, and reduce the stem-loop potential with prediction on the secondary structure of *alg-3/4* targeted small RNAs. By eliminating stem-loop structure at NYN-3 binding site and investigating whether *gfp* sRNAs still exists, we might be able to determine whether secondary structure is required for NYN-3 recognizing *alg-3/4* targeted mRNAs. Besides, this study only focused on one representative gene; more reporters need to be established to obtain global view of NYN-3 targeting rules.

Due to the transient interaction between ALG-3 and *alg-3/4* targeted mRNA template, we identified only eight candidate genes for new transcription start sites (TSSs). Given that previous studies mostly focus on hermaphrodites, there is limited information regarding TSSs specifically showed in males. To comprehensively understand the rules governing 26G small RNA biogenesis, it is crucial to identify all alternative TSSs of alg-3/4 targeted mRNAs in

doi:10.6342/NTU202403521

males. So far, we cannot conclude NYN-3 recognition of *alg-3/4* targeted mRNAs is based on TSSs for all *alg-3/4* targeted genes.

3-3 IFE-1 potentially guides NYN-3 recognition of *alg-3/4* targeted mRNA for the biogenesis of 26G small RNA

Our results indicate that IFE-1 would interact with NYN-3 at position nearby P-granule at the proximal germline, a crucial site where spermatogenesis is tightly regulated by various transcription factors and RNA regulators. Notably, the absence of either IFE-1 or NYN-3 leads to the absence of 26G small RNAs, resulting in complete sterility at 25 degrees Celsius. These observations strongly suggest that IFE-1 and NYN-3 may operate within the same pathway governing the biogenesis of 26G small RNAs. Thus, we proposed a model: when alg-3/4-targeted mRNAs are exported from the nucleus and localized to the P-granule, IFE-1 recognizes the cap structure and interacts with NYN-3. Subsequently, NYN-3 cleaves the mRNA sequence, generating a template for the RNA-dependent RNA polymerase RRF-3 to synthesize 26G small RNAs (Figure 5). These 26G small RNAs play an indispensable role in sperm function, particularly at 25 degrees Celsius.

However, the interdependency between IFE-1 and NYN-3 in recognizing alg-3/4 targeted mRNAs remains unclear. To understand the role of IFE-1 and NYN-3 in binding to *alg-3/4* targeted mRNAs, we can perform IFE-1or NYN-3 immunoprecipitation and confirm the enrichment of *alg-3/4* targeted mRNAs on one protein when the other is absent. Moreover, we can sequence the capped RNAs bound by IFE-1 and compare them with *alg-3/4* targeted

doi:10.6342/NTU202403521

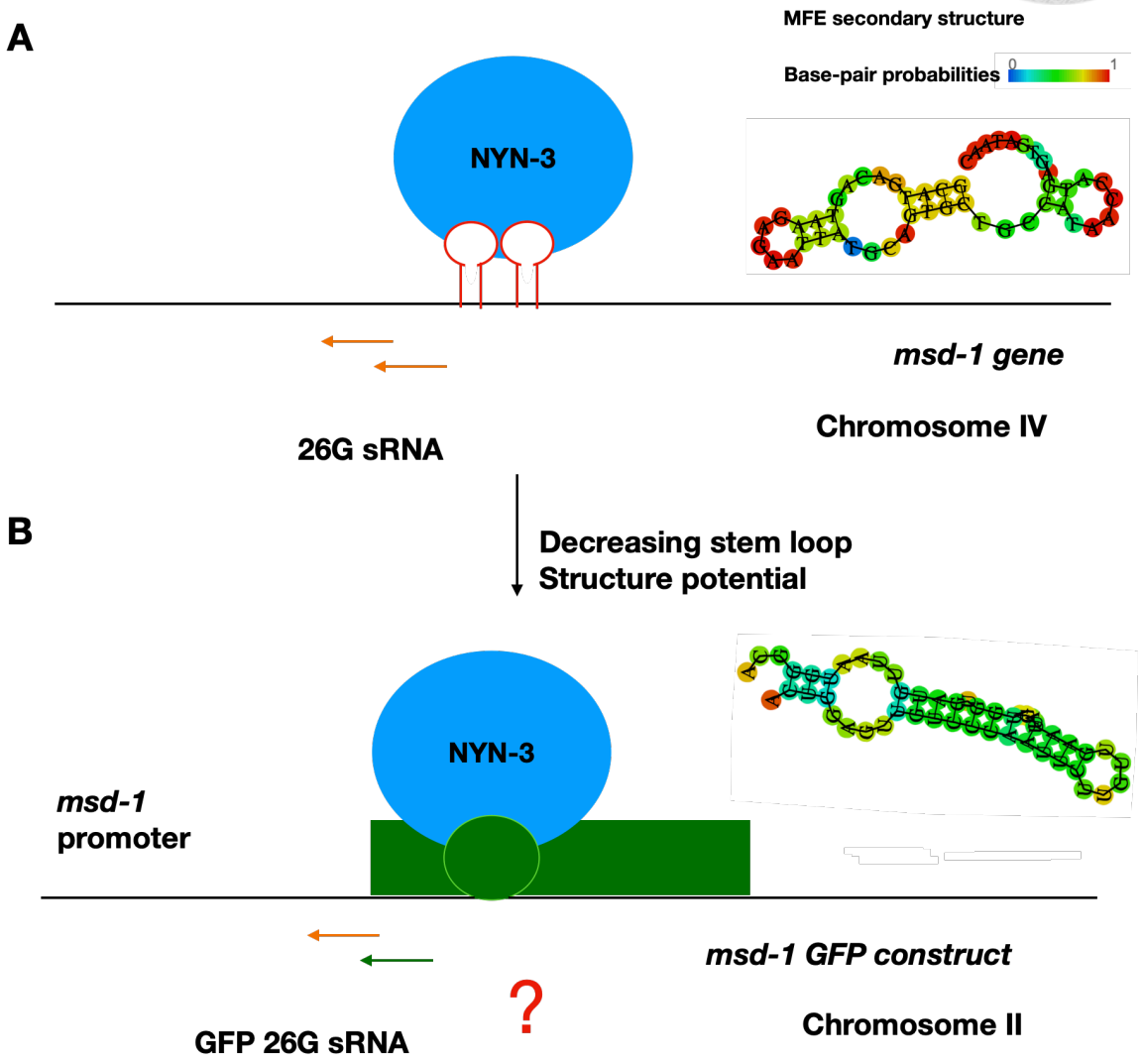
mRNAs to address whether all *alg-3/4* targeted mRNAs are associated with IFE-1 at their transcription start sites (TSSs).

All these studies may help us get a more comprehensive insights into the rules governing the biogenesis of 26G small RNAs and understand how these specific sRNAs are prepared for specific gene regulation during different developmental stages.

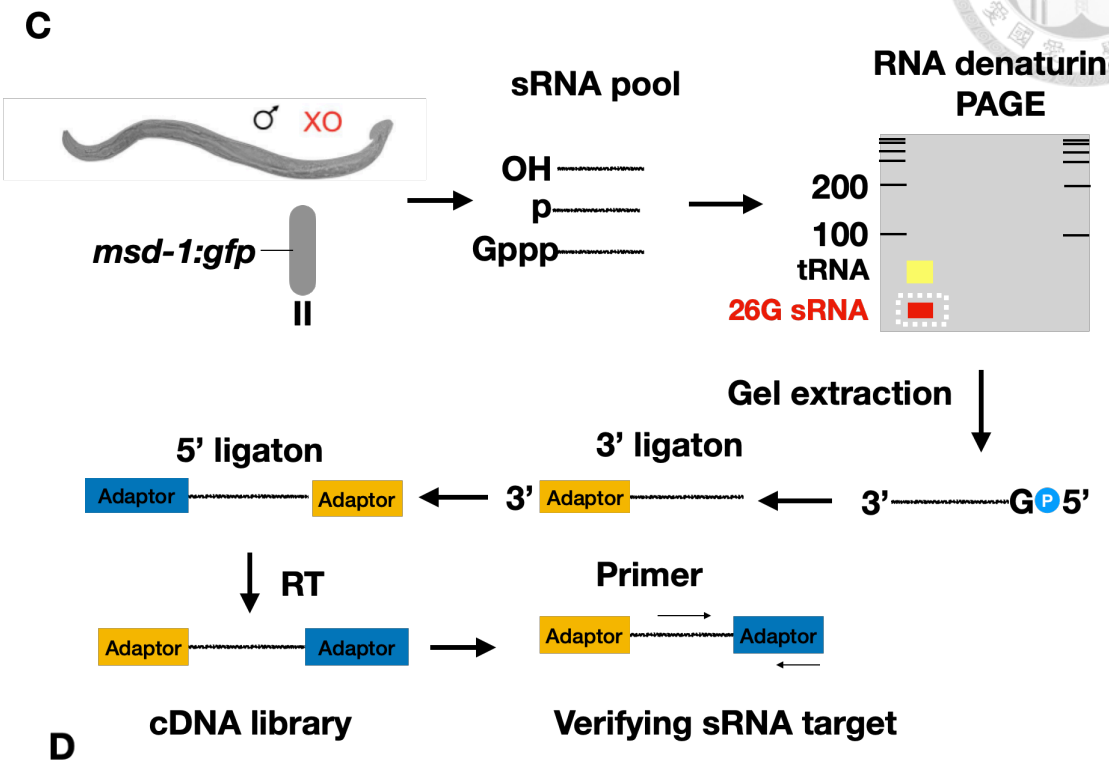
doi:10.6342/NTU202403521

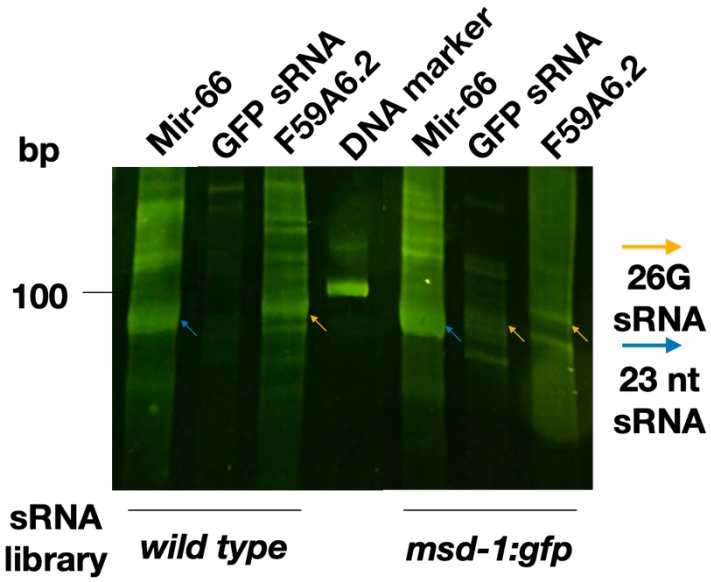
Chapter 5. Figures and Tables

Figure 1

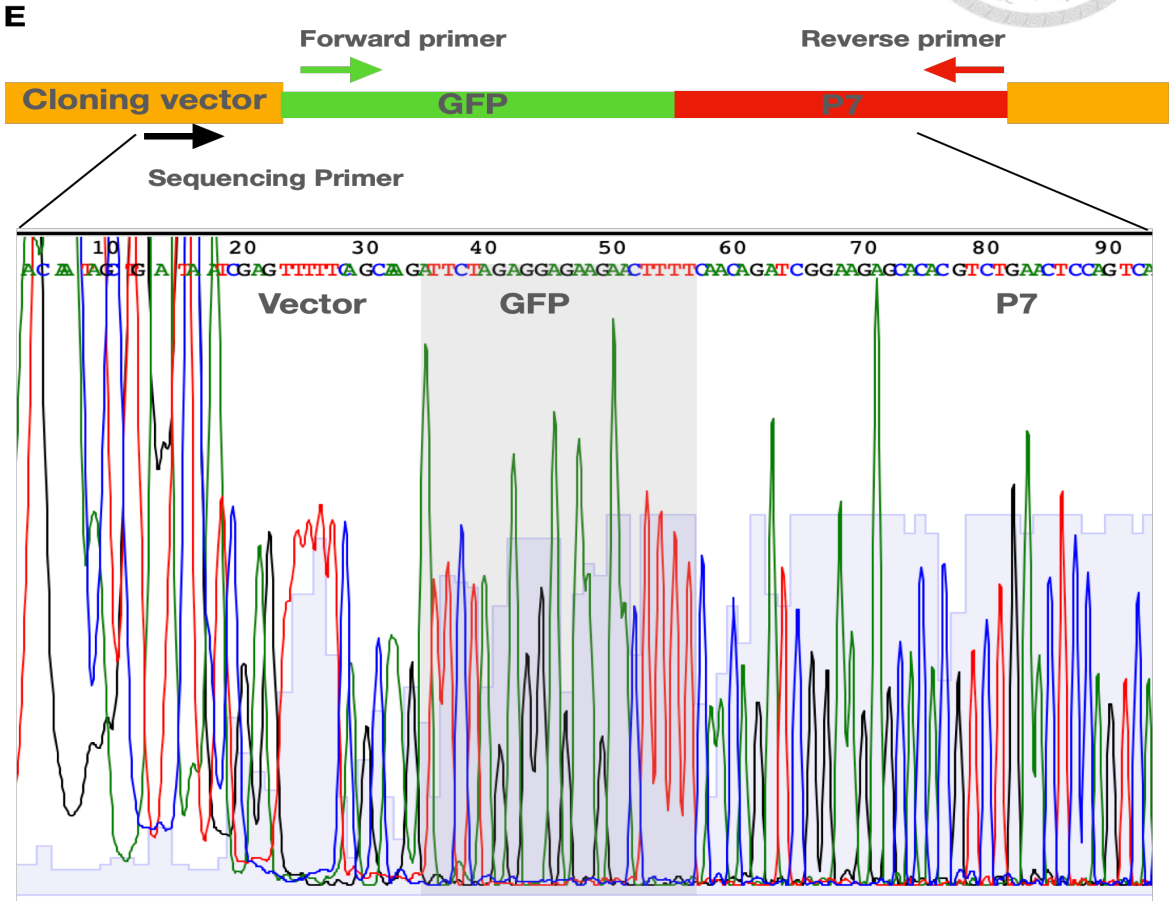


doi:10.6342/NTU202403521





doi:10.6342/NTU202403521



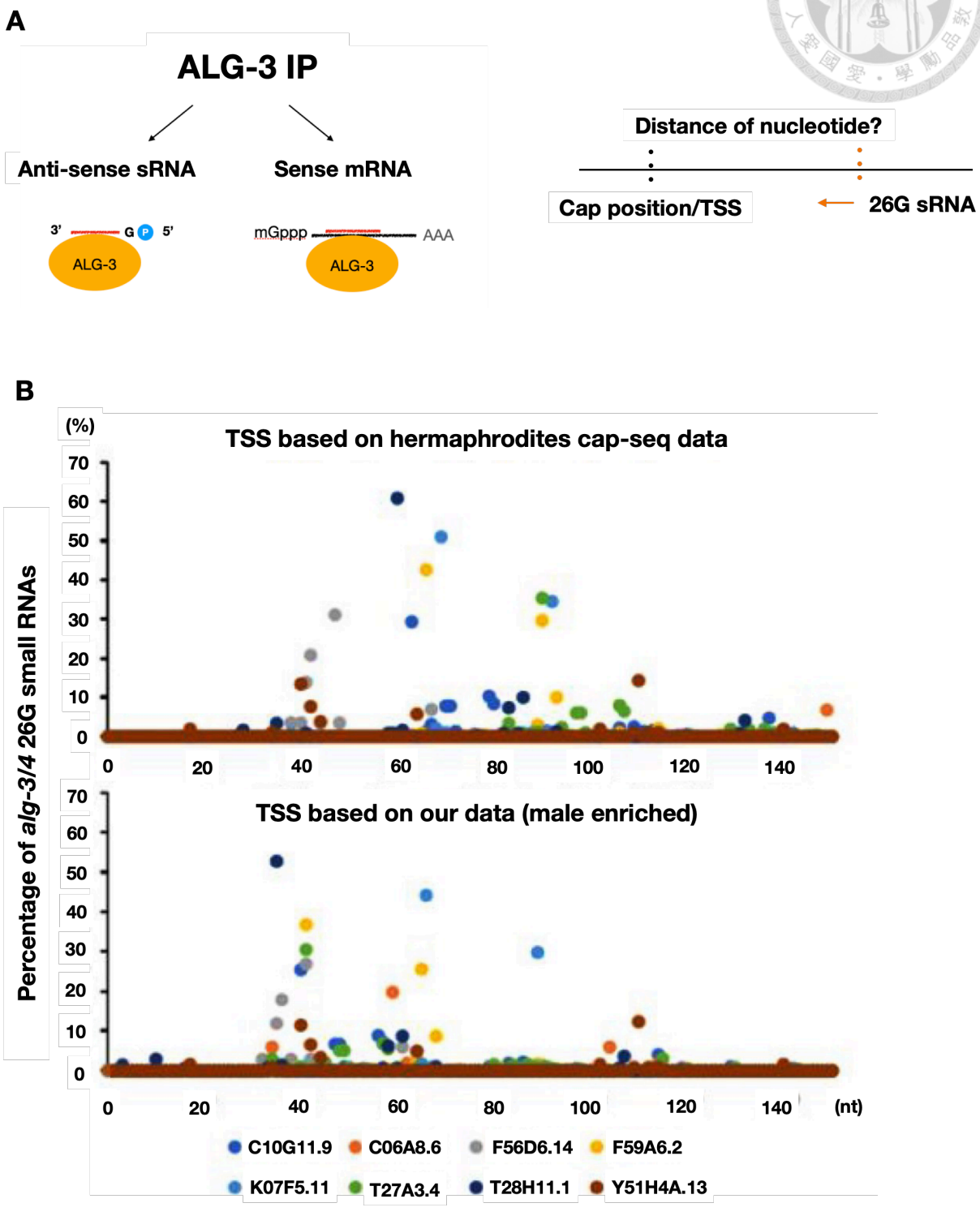
doi:10.6342/NTU202403521

Figure 1. The *gfp* small RNAs exist in *alg-3/4* small RNA reporter strain.

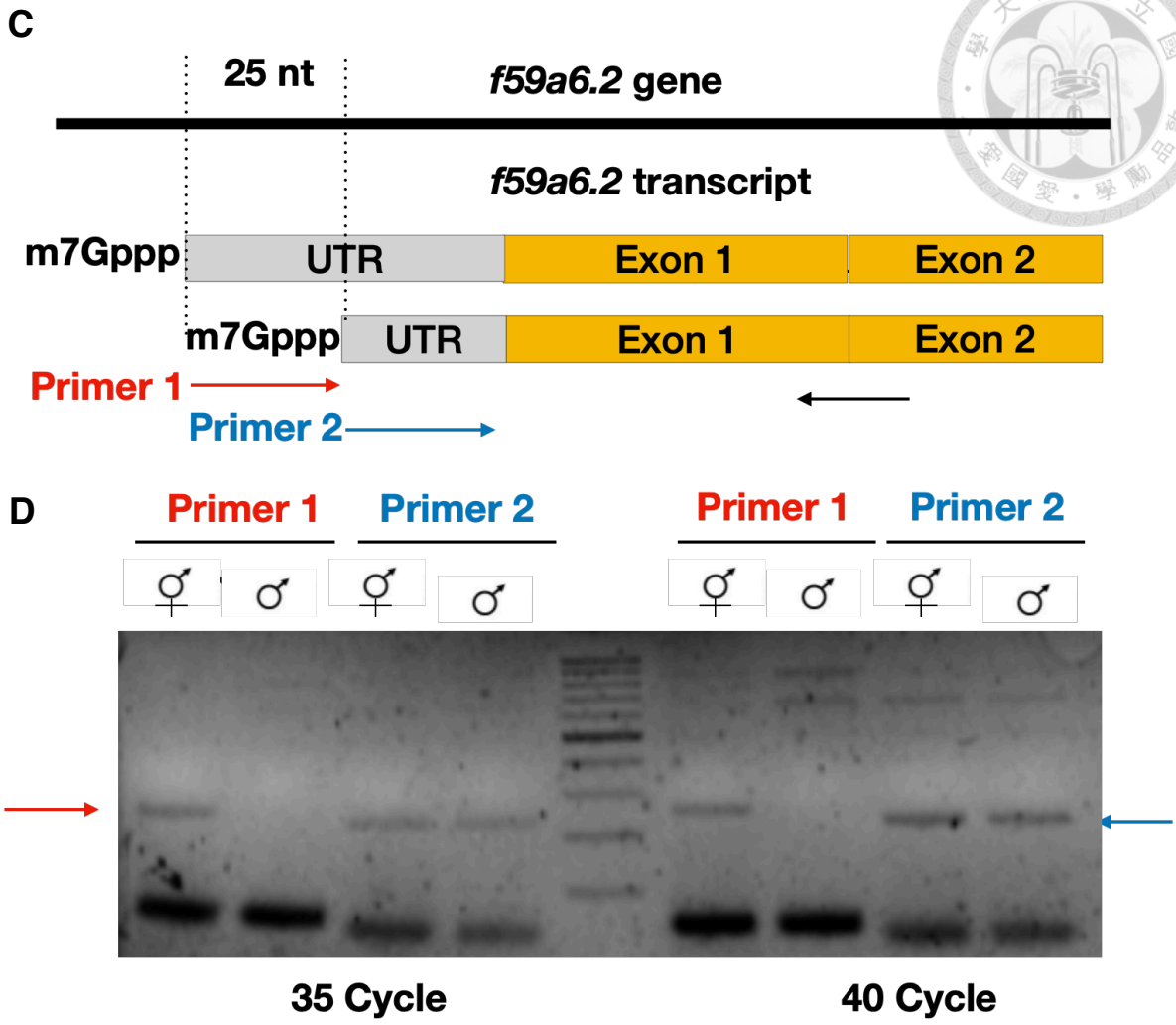
(A) *msd-1* gene is located on chromosome four and it is one of the targets of 26G sRNAs. NYN-3 recognize NYN-3 binding site and cleave the mRNA and generate RNA template for RNA-dependent RNA polymerase RRF3. RRF-3 then sequentially synthesize two 26G sRNAs. A predicted secondary structure of the sequence at NYN-3 binding site is show on the right, with base pairing probability indicated in different colors. (B) Construct of the *msd-1::gfp* reporter. In the reporter, chromosome two contains a single-copy insertion of the *alg-3/4* targeted gene, msd-1, fused with GFP sequence at the position where 26G sRNA is located and recognized by NYN-3. The GFP sequence alters the original msd-1 sequence, reducing the potential for stem loop structure. Whether NYN-3 can still recognize the modified mRNA and whether 26G small RNA corresponding to *gfp* sequencing can still be synthesized were further investigated. (C) Small RNA library preparation schematic. Small RNAs from male worms were first selected by size using denaturing polyacrylamide gel, and then underwent ligation with adaptors, followed by reverse transcription (RT). The presence of targeted small RNAs in the library can be confirmed using designed primers. (D) The presence of GFP 26G small RNAs was verified using RT-PCR and DNA-PAGE. *mir-66* and *f59a6.2* small RNAs served as positive controls in the final small RNA library. Primers used for detecting *gfp* 26G small RNA are shown in (E), indicated by red and green arrows. (E) Sanger sequencing of amplicons of *gfp* 26G small RNA. The amplicons of *gfp* 26G small RNA were ligated with cloning vector and transformed into bacteria for amplification. Cloning vectors containing *gfp* 26G small RNA sequence were sequenced using primer indicated by black arrow

doi:10.6342/NTU202403521

Figure 2



doi:10.6342/NTU202403521



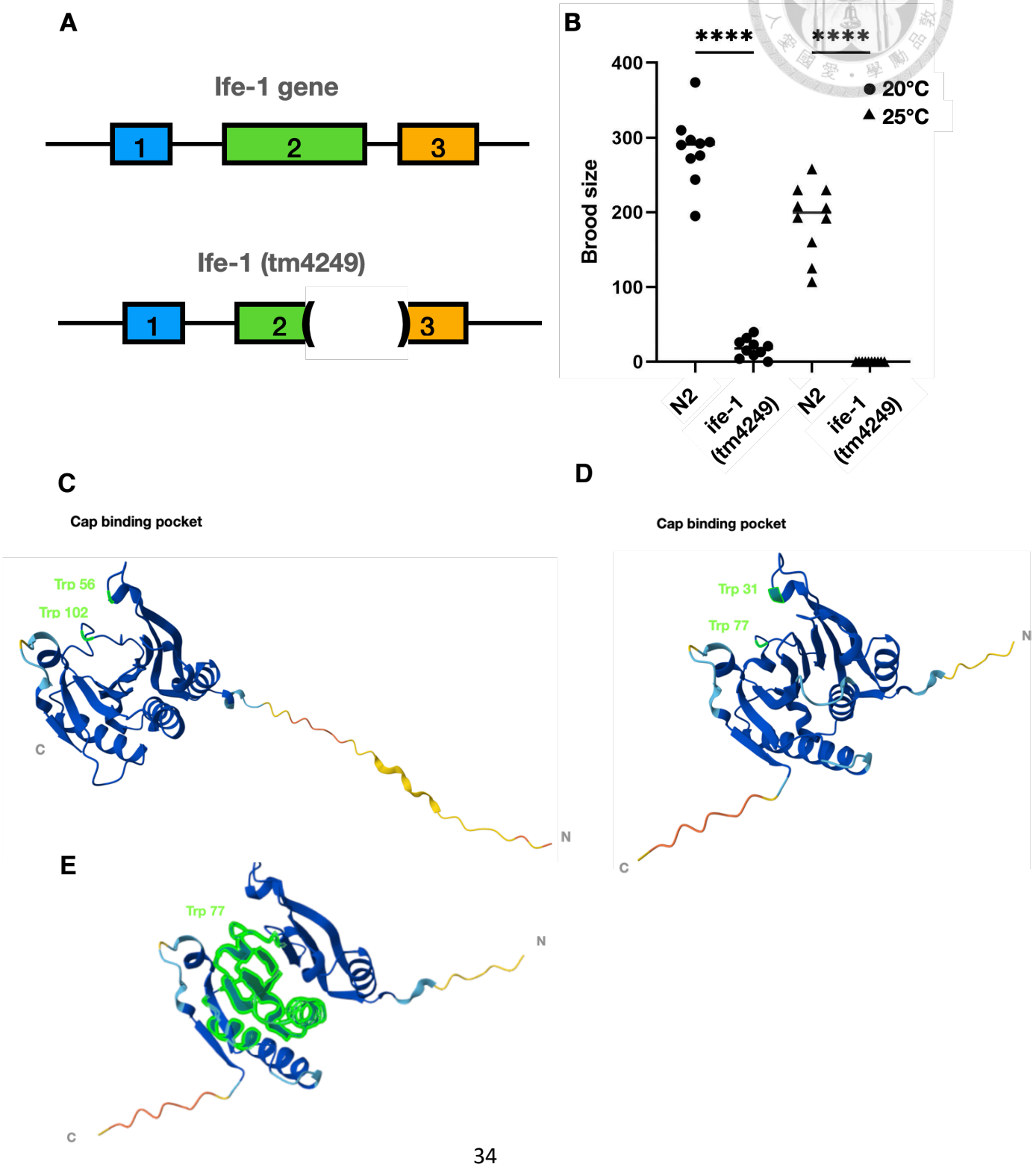
doi:10.6342/NTU202403521

Figure 2. **The discovery of alternative transcripts for *alg-3/4* targeted mRNAs.**

(A) In previous study, ALG-3 targeted mRNAs and sRNAs were immunoprecipitated for sequencing. The 5' end of 26G sRNAs was correlated with transcription start site/cap position and the distance between them was examinated. (B) Relative position of 5' end of 26G small RNAs and transcription start sites (TSSs). TSSs were identified from the adult cap-seq database (Top) or ALG-3-bound sense mRNA (Bottom). Eight candidate genes obtained from ALG-bound mRNAs sequencing were shown in different colors of dots. X axis indicates the distance between TSS and 5' end of 26G sRNAs. nt stands for nucleotides.Y axis indicates the percentage of *alg-3/4* 26G small RNAs. (C) Primers design for validating alternative transcripts of *f59a6.2*. Primer 1 designed based on the TSS identified from the ault cap-seq database, while primer 2 desigend based on the TSS identified in our studies. Revese primer includes exon-exon juction of Exon1 and Exon2. (D) Characterization of *f59a6.2* TSS from male and hermaphrodites using RT-PCR. Male worms are under *fog-2(q71)* background and hand-picked to ensure the alternative transcripts are biased in sexes.

doi:10.6342/NTU202403521

Figure 3



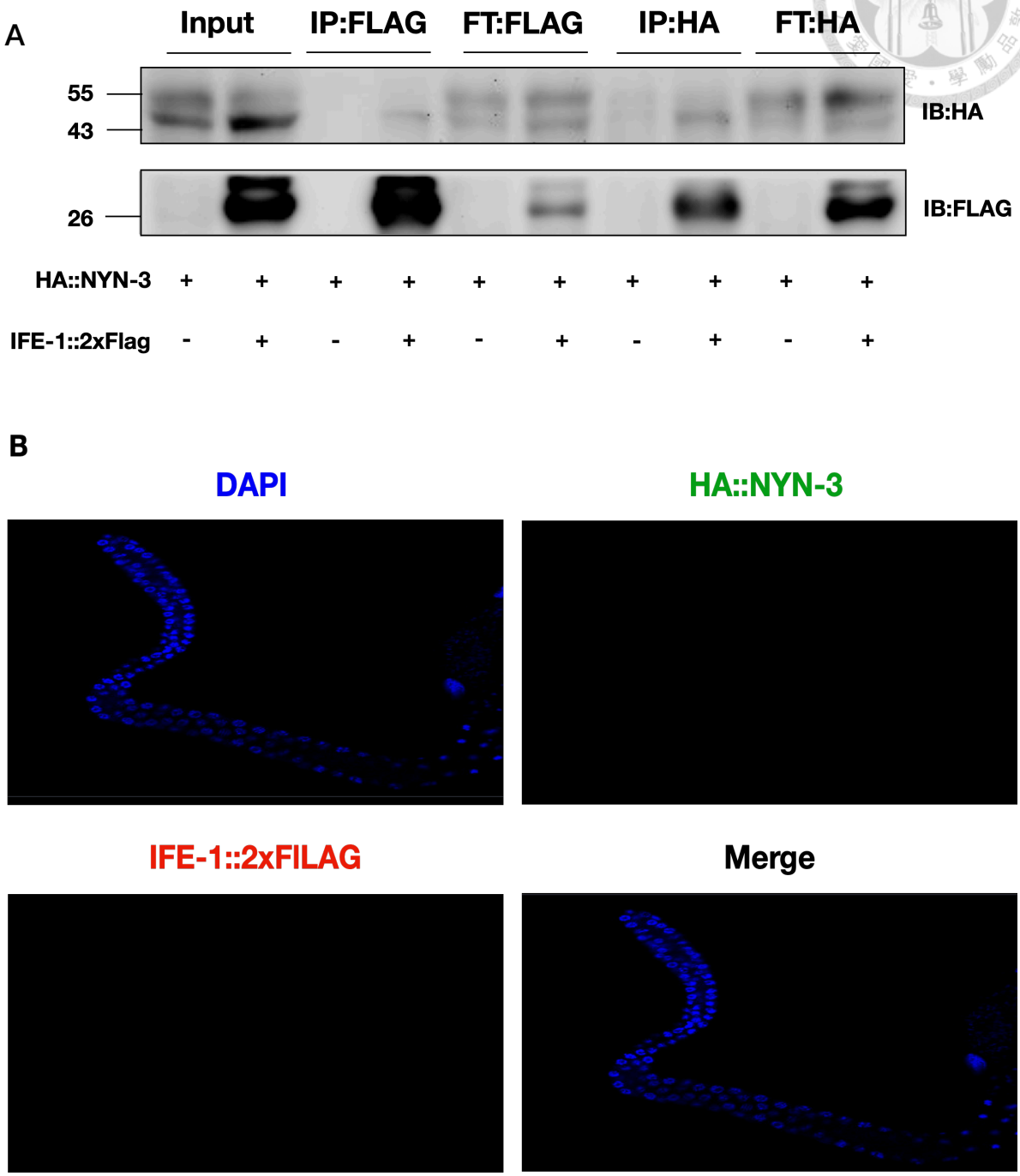
doi:10.6342/NTU202403521

Figure 3. Characterization of ife-1 wild type and ife-1 (tm4249) null mutant.

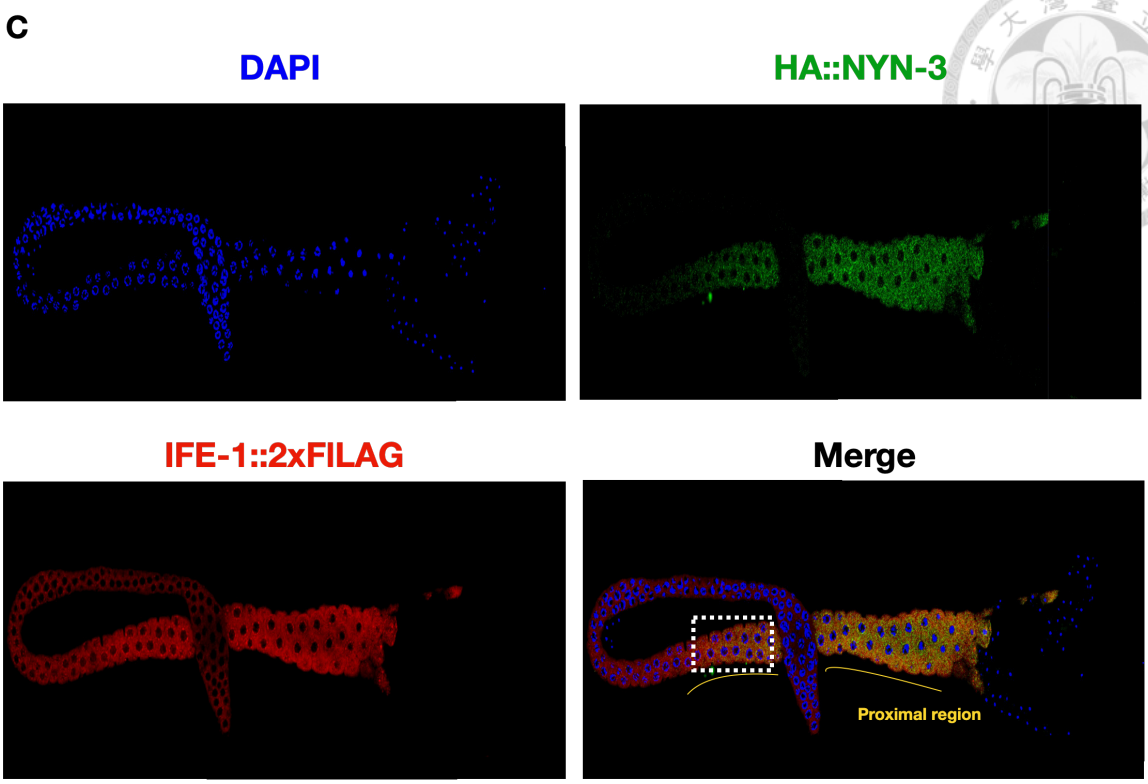
(A) Depiction of *ife-1* gene and locatoin of 343 bp tm4249 deletion. Exons are shown in colored boxes. (B) Brood size counts of Wild Type (N2) and IFE-1null mutant (*ife-1(tm4249)*) under 20℃ and 25℃. (N=10, ****P<0.0001, one-way ANOVA). (C-E) Representative images of the predicted protein structure of (C) human eukaryotic translation factor 4E (EIF4E) (D) IFE-1 (E) IFE-1(tm4249) null mutant. All the protein structures were predicted by AlphaFold 2. (C-D) Cap binding pockets containing conserved tryptophan residues are labeled in black, tryptophan (Trp) residues are marked in green, and C terminus, N terminus are labeled in gray. (E) For IFE-1null mutant, the protein residues affected by tm4249 deletion are labeled in green.

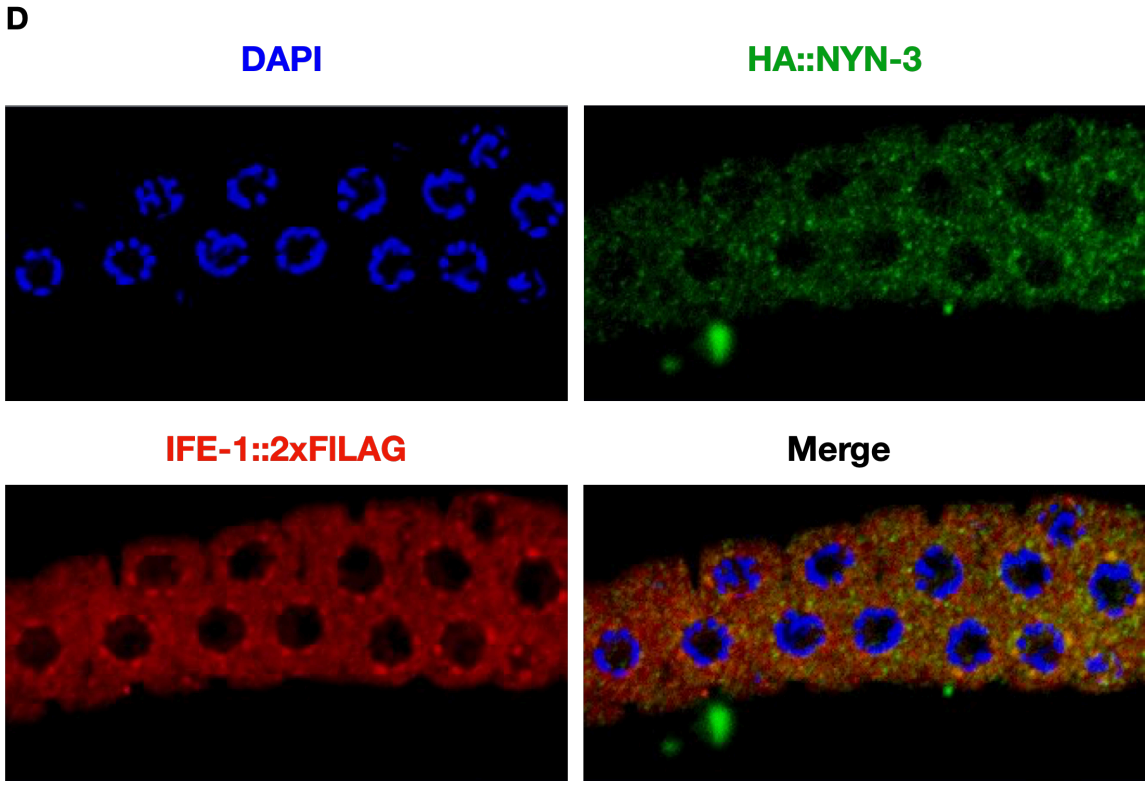
doi:10.6342/NTU202403521

Figure 4

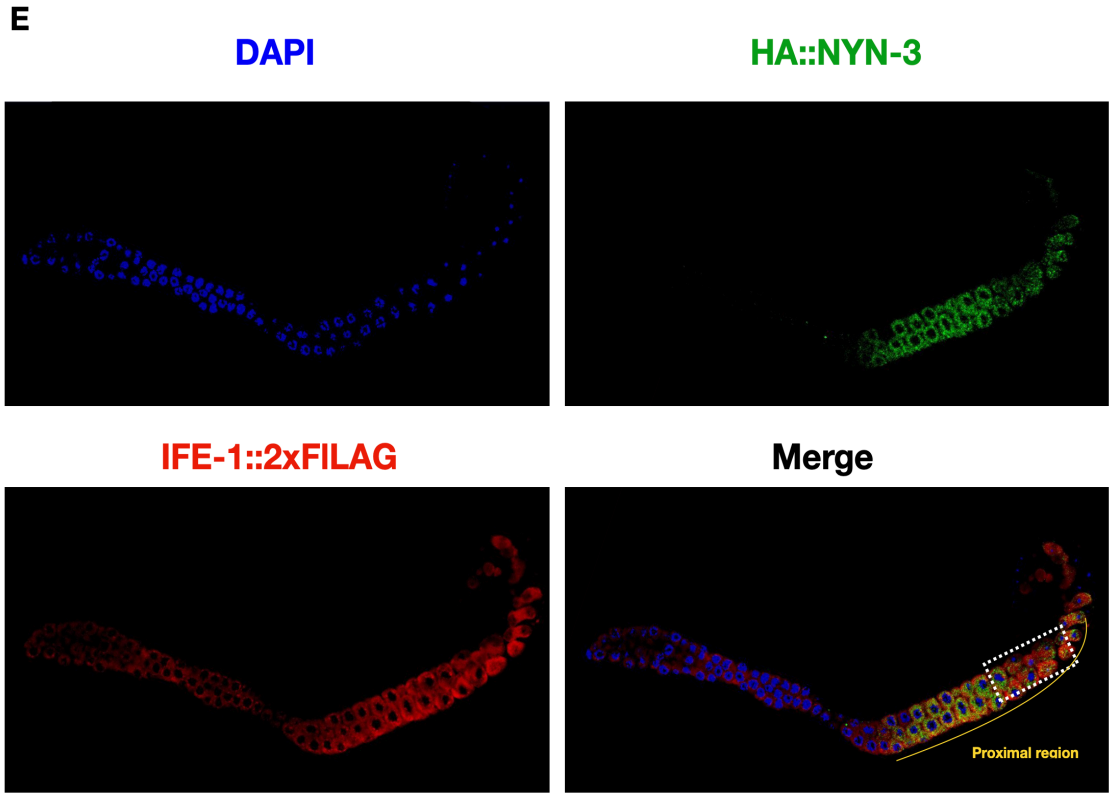


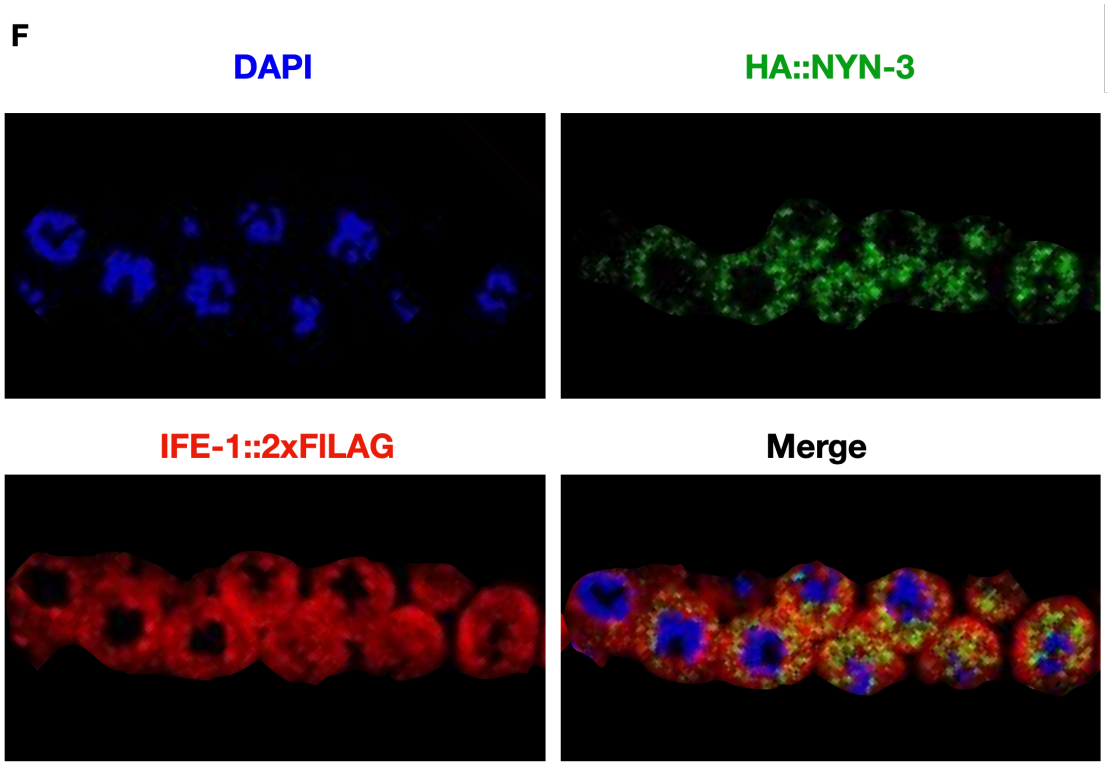
doi:10.6342/NTU202403521



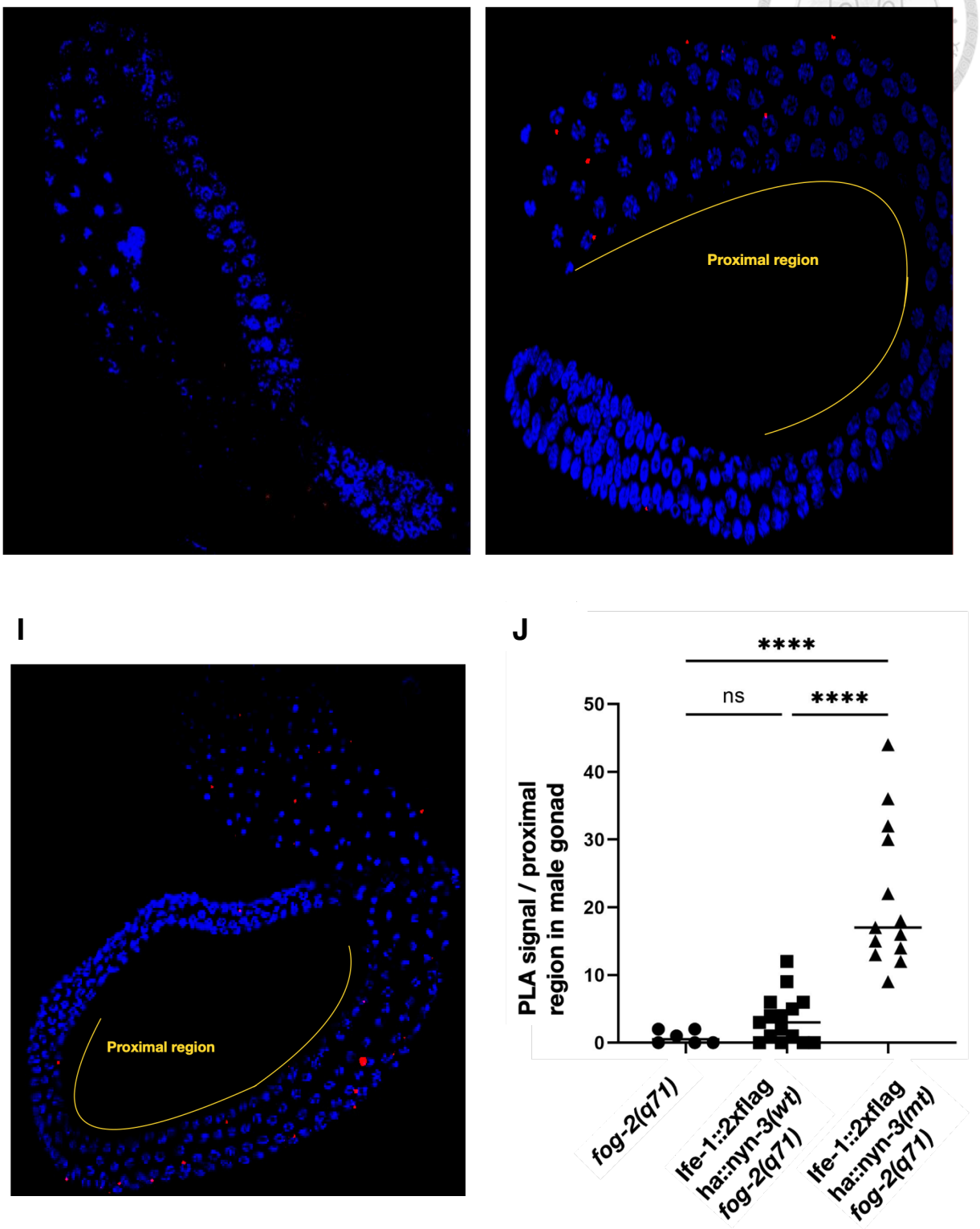


doi:10.6342/NTU202403521





doi:10.6342/NTU202403521



doi:10.6342/NTU202403521

Figure 4. Transient interaction between NYN-3 and IFE-1.

(A) Co-Immunoprecipitation analysis of NYN-3 and IFE-1 followed by western blotting. Total protein lysate was extracted from *ife-1::2x flag; ha::nyn-3(c29f5); fog-2(q71)* worms. 1mg, 7 mg of total protein lysate was used for Flag-immunoprecipitation (IP) and HA-IP, respectively. 90 μg of total protein lysate was used for input and flow through (FT) in western blot. (B-F) Immunostaining of IFE-1::2xFLAG and HA::NYN-3 in the male gonad of (B) wild type [*fog-2 (q71)*], (C) functional NYN-3 [*ife-1::2x flag; ha::nyn-3(c29f5); fog-2(q71)*] and (E) NYN-3 catalytic inactive mutant [*ife-1::2x flag; ha::nyn-3(ht-44); fog-2(q71)*]. (D)(F) The areas in dotted line box were enlarged to demonstrate potential interaction between IFE-1 and NYN-3. Wild-type fog-2 (q71) male worms served as experimental controls. DNA in the nucleus was stained with 4',6-diamidino-2-phenylindole (DAPI). DAPI, anti-HA and anti-FLAG staining are shown in the merged image. Proximal region of male gonad was labeled by yellow lines. (G-I) Proximity ligation assay (PLA) of (G) wild type [*fog-2 (q71)*], (H) functional NYN-3 [*ife-1::2x flag; ha::nyn-3(c29f5); fog-2(q71)*] and (I) NYN-3 catalytic inactive mutant [*ife-1::2x flag; ha::nyn-3(ht-44); fog-2(q71)*]. PLA signals (red) indicate the interaction between IFE-1 and NYN-3. DAPI signals (blue) show the nucleus DNA. (D) Quantification of PLA signals in the proximal region of male gonad. PLA signals were counted in the proximal region and distal region of male gonad in individual worm. Average PLA signals in the proximal region of male gonad were normalized by (total PLA signals in proximal region – total PLA signals in distal region)/ sample number (N), N=6 for *fog-2 (q71)*, N=15 for *ife-1::2x flag; ha::nyn-3(c29f5); fog-2(q71)*, N=13 foe *ife-1::2x flag; ha::nyn-3(ht-44); fog-2(q71)*. (****P<0.0001, one-way ANOVA).

doi:10.6342/NTU202403521

Figure 5

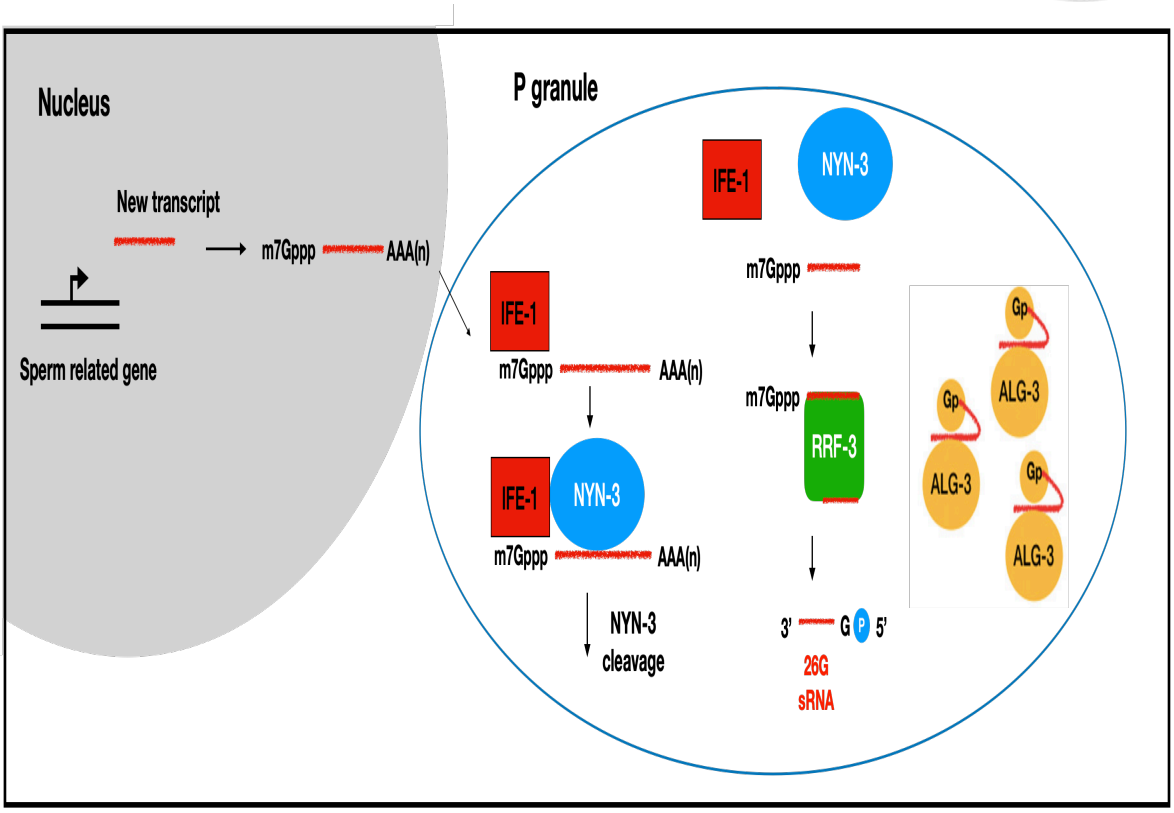


Figure 5. Proposed model of the role of IFE-1 in 26G sRNAs biogenesis.

doi:10.6342/NTU202403521

Table

Table-I Primers and adaptors used for small RNA library preparation.

3’ linker	5’-AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT/3pC3/ -3’
5’ linker	5’ - /5Sp9/CarGrArCrGrUrGrUrGrCrUrCrUrUrCrCrGrArUrCrU -3’
Primer for reverse transcription	5’-ACACTCTTTCCCTACACGACGCTCTTCCGATCT-3’
5’ adaptor (P7)	5’- CAAGCAGAAGACGGCATACGAGATGGAATCTCGTGACTGGAG TTCAGACGTGTGCTCTTC-3’
3’ adaptor (P5) *fog-2* library	5’- AATGATACGGCGACCACCGAGATCTACACTATAGCCTACACT CTTTCCCTACACGACGC-3’
3’ adaptor (P5) *msd-1::gfp* library	5’- AATGATACGGCGACCACCGAGATCTACACATAGAGGCACACT CTTTCCCTACACGACGC-3’

Table-II Primers for target small RNA detection.

mir-66	Forward 5’-TCACATCCCTAATCAGTGTCATG -3’ Reverse 5’-CAAGCGAAGACGGCATACGAGATGGAATCTCGTGAC TGGAGTTCAGACGTGTGCTCTTC-3’
f59a6.2 26G small RNA	Forward 5’- GCCACCACTGGAGTTCCTCCTTCTGC-3’ Reverse 5’-CAAGCGAAGACGGCATACGAGATGGAATCTCGTGAC TGGAGTTCAGACGTGTGCTCTTC-3’
gfp small RNA	Forward 5’-TCTAGAGGAGAAGAACTTTT-3’ Reverse 5’-CAAGCGAAGACGGCATACGAGATGGAATCTCGTGAC

doi:10.6342/NTU202403521

	TGGAGTTCAGACGTGTGCTCTTC-3’
Sanger sequencing cloning vector	5’- CGACTCACTATAGGGAGAGCGGC -3’

Table-III Primers for alternative transcripts detection.

f59a6.2 longer transcript based on TSSs identified previously	Forward 5’-AGTTTAAAAATTTCGTAGAATAATT-3’ Reverse 5’-AGTCCCTTCTTCGTGTGCAT-3’
f59a6.2 shorter transcript based on newly identified TSSs	Forward 5’-TATTCTGAATTTTTCGATTATCG-3’ Reverse 5’-AGTCCCTTCTTCGTGTGCAT-3’

doi:10.6342/NTU202403521

Reference

1. Song, J.-J., et al., *Crystal structure of Argonaute and its implications for RISC slicer activity.* science, 2004. **305**(5689): p. 1434-1437.
2. Hauptmann, J., et al., *Turning catalytically inactive human Argonaute proteins into active slicer enzymes.* Nature structural & molecular biology, 2013. **20**(7): p. 814-817.
3. Conine, C.C., et al., *Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans.* Proceedings of the National Academy of Sciences, 2010. **107**(8): p. 3588-3593.
4. Conine, C.C., et al., *Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. elegans.* Cell, 2013. **155**(7): p. 1532-1544.
5. Tsai, H.-Y., H.-T. Cheng, and Y.-T. Tsai, *Biogenesis of C. elegans spermatogenesis small RNAs is initiated by a zc3h12a-like ribonuclease.* Science Advances, 2022. **8**(32): p. eabm0699.
6. Grimson, A., et al., *Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals.* Nature, 2008. **455**(7217): p. 1193-1197.
7. Chu, D.S. and D.C. Shakes, *Spermatogenesis.* Germ cell development in C. elegans, 2013: p. 171-203.
8. McCarter, J., et al., *On the control of oocyte meiotic maturation and ovulation incaenorhabditis elegans.* Developmental biology, 1999. **205**(1): p. 111-128.
9. Yadav, R.P. and N. Kotaja, *Small RNAs in spermatogenesis.* Molecular and cellular endocrinology, 2014. **382**(1): p. 498-508.

doi:10.6342/NTU202403521

/10. Conine, C.C., et al., *Small RNAs gained during epididymal transit of sperm are essential for embryonic development in mice.* Developmental cell, 2018. **46**(4): p. 470-480. e3.

11. Ozata, D.M., et al., *PIWI-interacting RNAs: small RNAs with big functions.* Nature Reviews Genetics, 2019. **20**(2): p. 89-108.

12. Yamashiro, H. and M.C. Siomi, *PIWI-interacting RNA in Drosophila: biogenesis, transposon regulation, and beyond.* Chemical reviews, 2017. **118**(8): p. 4404-4421.

13. Batista, P.J., et al., *PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans.* Molecular cell, 2008. **31**(1): p. 67-78.

14. Tóth, K.F., et al., *The piRNA pathway guards the germline genome against transposable elements.* Non-coding RNA and the Reproductive System, 2016: p. 51-77.

15. Belicard, T., P. Jareosettasin, and P. Sarkies, *The piRNA pathway responds to environmental signals to establish intergenerational adaptation to stress.* BMC biology, 2018. **16**: p. 1-14.

16. Sassone-Corsi, P., *Unique chromatin remodeling and transcriptional regulation in spermatogenesis.* Science, 2002. **296**(5576): p. 2176-2178.

17. Dai, P., et al., *A translation-activating function of MIWI/piRNA during mouse spermiogenesis.* Cell, 2019. **179**(7): p. 1566-1581. e16.

doi:10.6342/NTU202403521

18. Vasale, J.J., et al., *Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway.* Proceedings of the National Academy of Sciences, 2010. **107**(8): p. 3582-3587.

19. Han, T., et al., *26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans.* Proceedings of the National Academy of Sciences, 2009. **106**(44): p. 18674-18679.

20. Gingras, A.-C., B. Raught, and N. Sonenberg, *eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation.* Annual review of biochemistry, 1999. **68**(1): p. 913-963.

21. Huggins, H.P., et al., *Distinct roles of two eIF4E isoforms in the germline of Caenorhabditis elegans.* Journal of cell science, 2020. **133**(6): p. jcs237990.

22. Amiri, A., et al., *An isoform of eIF4E is a component of germ granules and is required for spermatogenesis in C. elegans.* 2001.

23. Gajjar, G., et al., *Two germ granule eIF4E isoforms reside in different mRNPs to hand off C elegans mRNAs from translational repression to activation.* bioRxiv, 2024.

24. Henderson, M.A., et al., *A germline-specific isoform of eIF4E (IFE-1) is required for efficient translation of stored mRNAs and maturation of both oocytes and sperm.* Journal of cell science, 2009. **122**(10): p. 1529-1539.

25. Brenner, S., *The genetics of Caenorhabditis elegans.* Genetics, 1974. **77**(1): p. 71-94.

doi:10.6342/NTU202403521

26. Van Nostrand, E.L., et al., *Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP).* Nature methods, 2016. **13**(6): p. 508-514.

27. Gu, W., et al., *CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors.* Cell, 2012. **151**(7): p. 1488-1500.

28. Phillips, C.M., K.L. McDonald, and A.F. Dernburg, *Cytological analysis of meiosis in Caenorhabditis elegans.* Meiosis: Volume 2, Cytological Methods, 2009: p. 171-195.

29. Morrison, K.N., et al., *MFP1/MSD-1 and MFP2/NSPH-2 co-localize with MSP during C. elegans spermatogenesis.* microPublication biology, 2021. **2021**.

30. Karaki, S., et al., *The eukaryotic translation initiation factor 4E (eIF4E) as a therapeutic target for cancer.* Advances in Protein Chemistry and Structural Biology, 2015. **101**: p. 1-26.

31. Pugacheva, E.M., et al., *BORIS/CTCFL epigenetically reprograms clustered CTCF binding sites into alternative transcriptional start sites.* Genome Biology, 2024. **25**(1): p. 1-40.

32. Heger, P., B. Marin, and E. Schierenberg, *Loss of the insulator protein CTCF during nematode evolution.* BMC molecular biology, 2009. **10**: p. 1-14.

33. del Castillo-Olivares, A., M. Kulkarni, and H.E. Smith, *Regulation of sperm gene expression by the GATA factor ELT-1.* Developmental biology, 2009. **333**(2): p. 397-408.

doi:10.6342/NTU202403521