m6A is required for resolving progenitor identity during planarian stem cell differentiation

Yael Dagan[1]*, Yarden Yesharim[1]*, Ashley R. Bonneau[2,3,4], Tamar Frankovits[1], Schraga Schwartz[5], Peter W. Reddien[2,3,4], Omri Wurtzel[1,6,7]

[1] School of Neurobiology, Biochemistry, and Biophysics, The George S. Wise Faculty of Life Sciences, Tel Aviv University, 69978 Tel Aviv, Israel

[2] Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA

[3] Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[4] Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA

[5] Department of Molecular Genetics, Weizmann Institute of Science, 7610001 Rehovot, Israel

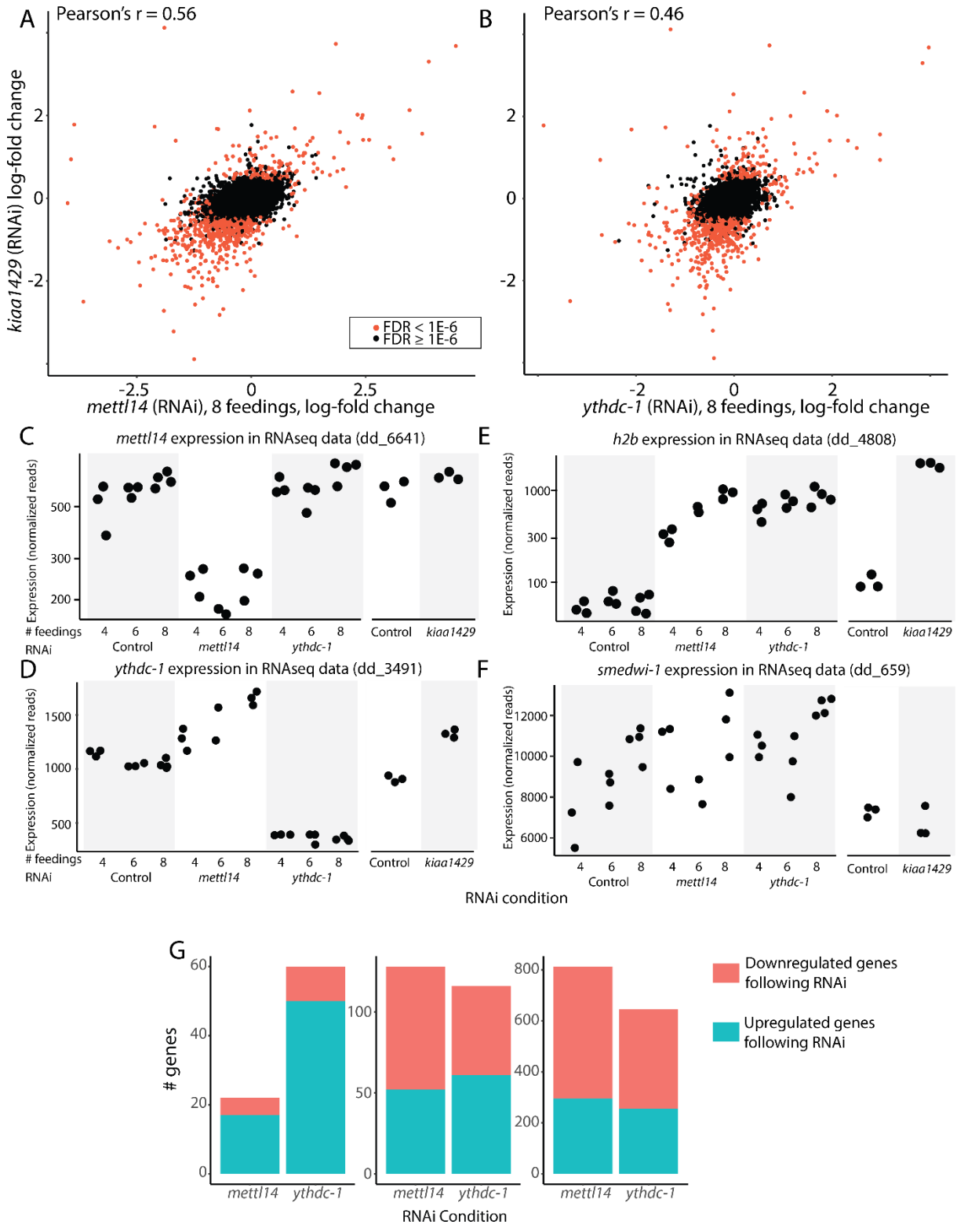
[6] Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel

* These authors contributed equally

[7] Correspondence: owurtzel@tauex.tau.ac.il

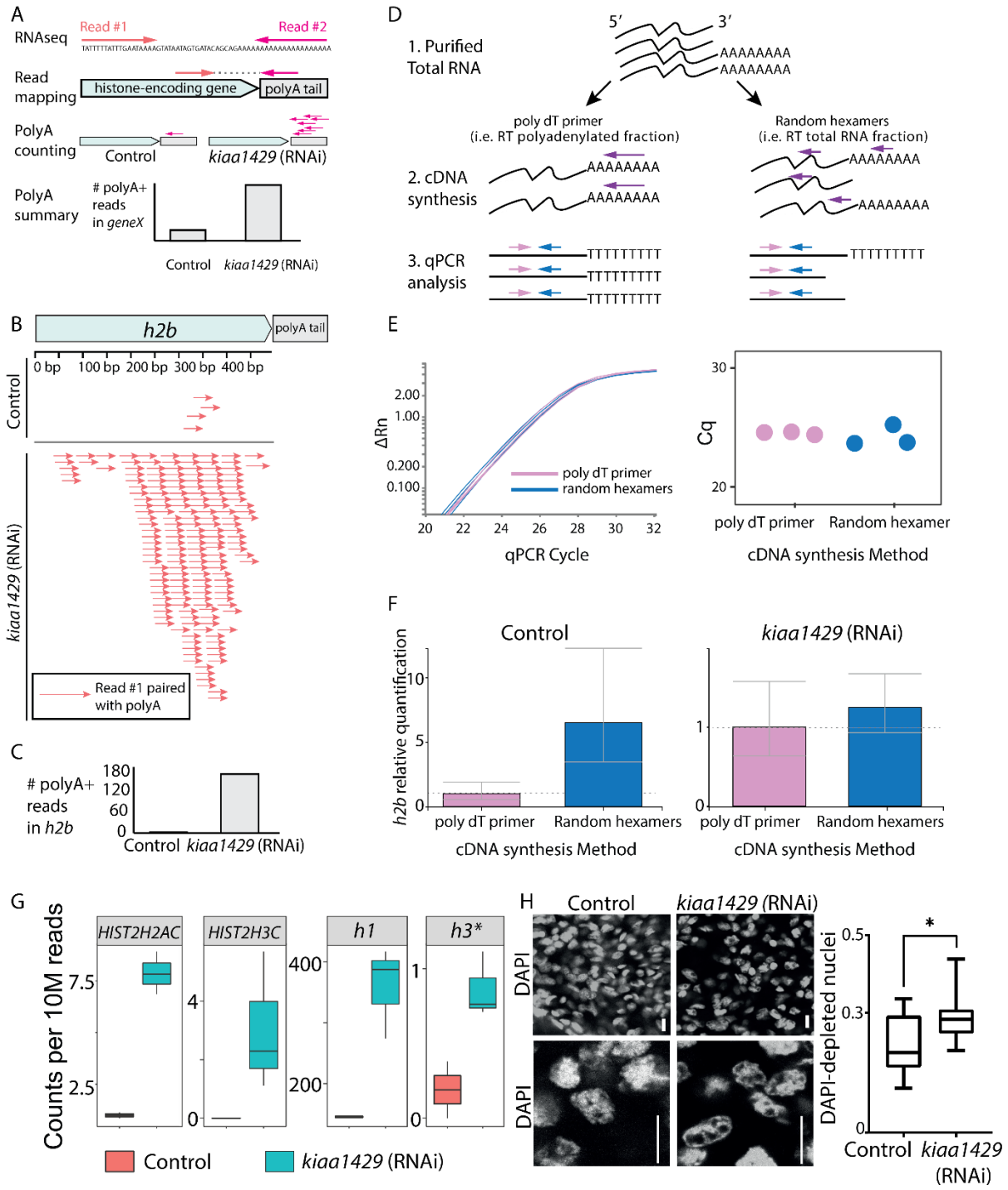
Table of contents

Appendix Figure S1



Appendix Figure S1. Specificity of RNAi of m6A genes. (A-B) Shown is the correlation in gene expression changes between experimental conditions. Each dot represents the log fold-change of a single gene, colored based on significance threshold (FDR threshold < 1E-6). For both panels the p-value of the correlation < 2.2E-16. (C-F) Shown is the normalized gene expression in RPKM for genes across the different libraries. RNAseq analysis validated the efficiency and the specificity of the RNAi (C-D). The expression of *h2b* was upregulated in all of the tested conditions (E). By contrast, the expression of *smedwi-1* did not change significantly following inhibition of m6A genes (F). (E) The number of significantly upregulated and downregulated genes is shown at each of the tested time points (FDR of genes included < 1E-5; Dataset EV2). In the early time point (four RNAi feedings) most of the genes were upregulated. However, at later time points, most differentially expressed genes are downregulated, which likely represent indirect effects of the RNAi, such as depletion of cell populations.

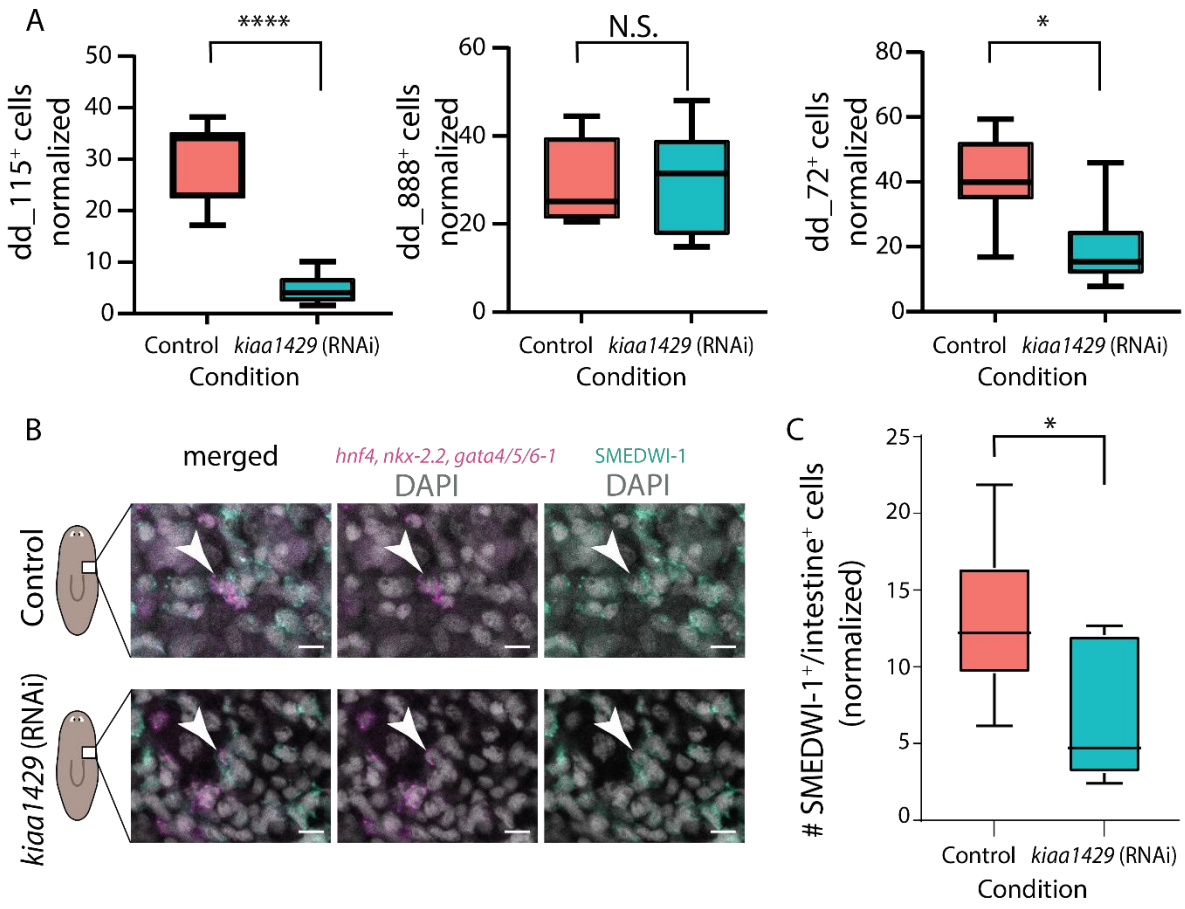
Appendix Figure S2



Appendix Figure S2. Analysis of histone transcript polyadenylation following inhibition of m6A genes. (A) Paired-end RNAseq data was used for identifying polyA containing histone transcripts. Reads mapped (read #1) to histone

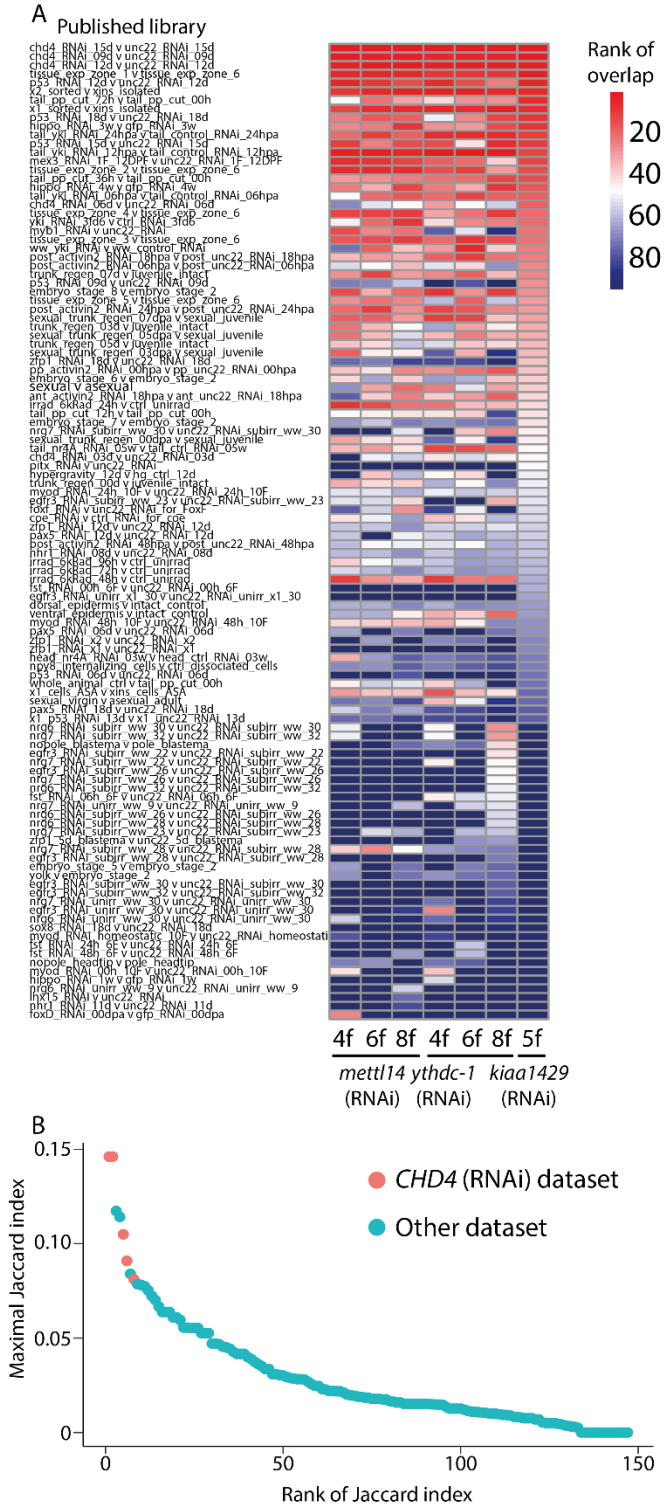
genes without proper read-pairing were computationally isolated, and their read pair (read #2) was collected from the raw data files. If the unmapped read (read #2) contained a poly-dT sequence, which is the reverse-complement of polyA, then the read was counted as a polyadenylated histone transcript. (B-C) Shown are RNAseq reads mapped to *h2b* (dd_4808; gene model SMESG000067906), which had a paired-read (read #2) containing polyA. Control sample (top) has fewer reads mapped to *h2b* compared to the *kiaa1429* (RNAi) sample (bottom). The total number of polyA containing read-pairs mapped to *h2b* is shown in panel C. (D) qPCR strategy for estimating the fraction of polyadenylated and non-polyadenylated *h2b*. Purified RNA (1) was reverse-transcribed with a poly-dT, or random hexamers (2). Then, qPCR is performed with internal primers (3). (E) The expression of *gapdh* was measured using RNA that was purified from wild type planarians. The analysis was used to estimate the efficiency of qPCR using cDNA produced either with poly-dT priming or using random hexamers. The change in reporter signal (y-axis) as a function of the qPCR cycle was similar for both groups (left), as is the quantification cycle (Cq, right). The analysis demonstrated that reverse-transcription with poly-dT and random hexamers was similarly efficient. (F) Shown in the relative quantification of polyadenylated *h2b* and non-polyadenylated *h2b* in control (left) and following *kiaa1429* (RNAi). The increase in polyadenylated *h2b* expression following *kiaa1429* (RNAi) was observed despite a lack of increase in total *h2b* expression (right). Error bars are the 95% confidence interval. (G) Shown is the number of polyA containing reads for several of the overexpressed histone components encoding genes in RNAseq data per 10M library reads (*h3 is the putative *histone 3* transcript that is transcribed from contig dd_25629, gene model SMESG000013634). Horizontal line indicates the median. (H) Comparison of DNA density by using DAPI-labeling on control and *kiaa1429* (RNAi) animals. DAPI-labeling, which correlates with chromatin density, showed an overabundance of DAPI-poor nuclei in *kiaa1429* (RNAi) animals. This suggested that *kiaa1429* (RNAi) chromatin packaging was altered, and indicated that euchromatin might be more prevalent in *kiaa1429* (RNAi), which might therefore affect transcriptional regulation. High magnification images (bottom) were acquired from different field-of-views from similar anatomical areas. Whiskers indicate the range of DAPI-poor-nuclei to nuclei ratio in the images, boxes indicate the interquartile range (IQR), and central band indicates the mean (n = 10 biological replicates per group). Student's t-test * $p < 0.05$. Scale = 10 µm.**

Appendix Figure S3



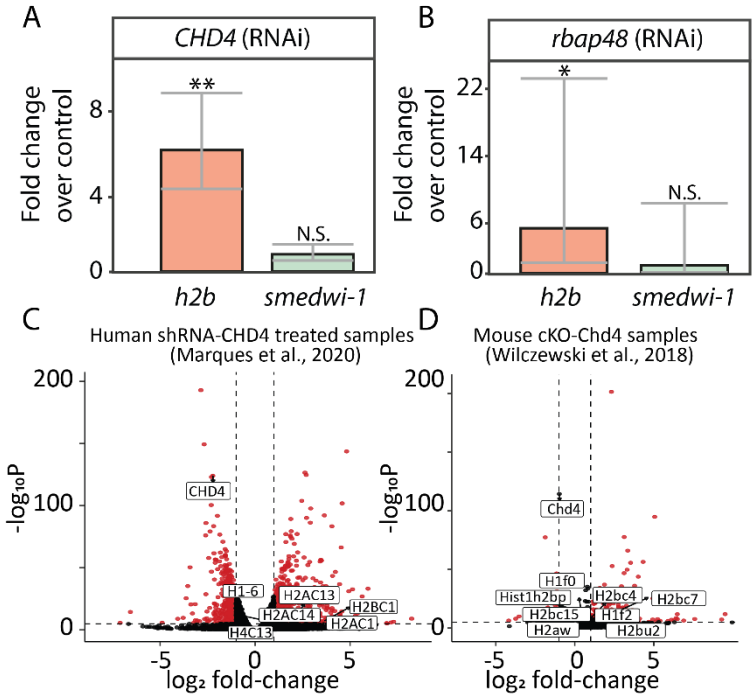
Appendix Figure S3. Reduction in intestinal cells following *kiaa1429* (RNAi). (A) Shown are cell counting of FISH images in *kiaa1429* (RNAi) and in control animals normalized by the area of counting (rectangular region anterior to the pharynx and posterior to the brain; Methods). Representative images are shown in Figure 6E (Two sided Student's t-test p = 1.2E-7 and 0.001 for dd_115 and dd_72, respectively followed by Bonferroni's correction). (B) Shown is co-labeling of recently integrated intestinal cells into to intestine structure by using a mix probes for detection of intestine-specific transcription factors (i.e., *hnf4, nkx-2.2,* and *gata4/5/6-1*), and anti-SMEDWI-1, which labels neoblasts and recent progeny (Guo *et al*, 2006). White arrowheads indicate double positive cell. Scale = 10 µm. (C) Quantification of co-labeled SMEDWI-1+/intestine+ cells is shown in Control and *kiaa1429* (RNAi) animals, based on cell counting from z-stack images (two sided Student's t-test = 0.02).

Appendix Figure S4



Appendix Figure S4. Overlap in differentially expressed genes. (A) Shown is the rank of overlap between the RNAseq libraries produced here and the published planarian datasets. The rank was determined by counting the number of overlapping and non-overlapping differentially expressed genes in the datasets produced here and in published datasets (FDR < 1E-5 and TPM > 10), and calculating the hypergeometric p-value for the overlap (Table EV1; blue to red, low and high ranked significance of overlap, respectively). (B) Plot showing the maximal Jaccard index calculated between up-to the top 200 differentially expressed genes in the published planarian datasets used here (Table EV1) and up-to the top 200 differentially expressed genes of each of the datasets produced here (Dataset EV2). Each dot represents a single dataset. *CHD4* (RNAi) datasets are highlighted in red.

Appendix Figure S5



Appendix Figure S5. *CHD4* (RNAi) recapitulates molecularly the inhibition of m6A genes. (A-B) Shown is a qPCR quantification of *h2b* and *smedwi-1* expression following RNAi of the NuRD component encoding genes *CHD4* (A) or *rbap48* (B) in comparison to control libraries (error bars indicate the 95% confidence interval; two-sided t-test Bonferroni corrected; ** - p < 0.01, * - p < 0.05). (C-D) volcano plot of human (C) and mouse (D) gene expression changes following inhibition or conditional knockout of CHD4 encoding gene (Marques *et al*, 2020; Wilczewski *et al*, 2018). Highlighted and labeled are significantly overexpressed histone-encoding genes, as well as the human or mouse *CHD4* homolog (red and black, significant and non-significant change in gene expression; adjusted p-value < 0.0001).

Appendix Table S1. Gene models of contig IDs in figures

Figure	Panel	Label in figure	Gene Model
Figure 2	A	dd_254	SMESG000073667
Figure 4	A	dd_783	SMESG000016810.1
Figure 4	A	dd_1837	SMESG000016779.1; SMESG000016781.1
Figure 4	A	dd_11930	SMESG000016771.1
Figure 4	A	dd_18852	SMESG000016780.1
Figure 4	A	dd_1616	SMESG000022742.1
Figure 4	A	dd_19404	NA
Figure 4	A	dd_15204	SMESG000081110.1
Figure 4	A	dd_20482	NA
Figure 4	A	dd_2089	SMESG000016811.1
Figure 4	A	dd_7261	NA
Figure 4	A	dd_2262	SMESG000038305.1
Figure 4	A	dd_1785	SMESG000005903.1; SMESG000006038.1
Figure 4	A	dd_1938	SMESG000007684.1
Figure 4	A	dd_4923	SMESG000063585.1
Figure 4	A	dd_6637	SMESG000071518.1
Figure 4	A	dd_2759	SMESG000081460.1
Figure 4	A	dd_3377	SMESG000026665.1
Figure 4	E	dd_175	SMESG000049722.1
Figure 4	E	dd_4575	SMESG000068883.1
Figure 4	E	dd_3194	SMESG000053627.1
Figure 5	A	dd_25629	SMESG000013634.1
Figure 5	H	dd_1837	SMESG000016779.1; SMESG000016781.1
Figure 5	H	dd_585	SMESG000062836.1
Figure 6	E	dd_72	NA
Figure 6	E	dd_888	SMESG000073691.1
Figure 6	E	dd_75	SMESG000027140.1
Figure 6	E	dd_115	SMESG000066413.1; SMESG000066416.1
Figure 7	E	dd_25269	SMESG000013634.1
Figure EV1	A	dd_6450	SMESG000034548.1
Figure EV1	A	dd_6641	SMESG000073528.1
Figure EV1	A	dd_4676	SMESG000074445.1; SMESG000074456.1; SMESG000074577.1; SMESG000074578.1
Figure EV1	A	dd_2322	SMESG000043489.1; SMESG000043668.1
Figure EV1	A	dd_7426	SMESG000019097.1
Figure EV1	A	dd_7282	SMESG000001341.1
Figure EV1	B	dd_3491	SMESG000018260.1
Figure EV1	B	dd_7891	SMESG000008619.1

Figure EV1	B	dd_8450	SMESG000003195.1
Figure EV1	B	dd_5578	SMESG000001953.1
Figure EV1	B	dd_7162	SMESG000039495.1
Figure EV1	B	dd_5316	SMESG000075874.1
Figure EV4	E	dd_3194	SMESG000053627.1
Figure EV4	G	dd_3194	SMESG000053627.1
Appendix Figure S1	A	dd_6641	SMESG000073528.1
Appendix Figure S1	B	dd_3491	SMESG000018260.1
Appendix Figure S1	C	dd_4808	SMESG000067906.1
Appendix Figure S1	D	dd_659	SMESG000036375.1
Appendix Figure S4	A	dd_115	SMESG000066413.1; SMESG000066416.1
Appendix Figure S5	A	dd_888	SMESG000073691.1
Appendix Figure S6	A	dd_72	NA

Appendix Table S2. Genes cloned in this study

Contig	Gene model	Primer ID	Sequence
dd_Smed_v4_4676_1	SMESG000074445	Kiaa1429_F	CCGCTATCCGTTGTTATTATGCG
dd_Smed_v4_4676_1	SMESG000074445	Kiaa1429_R	TCCGTAATCGTGGCCAGC
dd_Smed_v4_7282_1	SMESG000001341	rbm15_F	GGGAGTTATCTGATGGTCAAAGA
dd_Smed_v4_7282_1	SMESG000001341	rbm15_R	CCGGCACCGCTAACAGAA
dd_Smed_v4_6450_0_1	SMESG000034548	mettl3_F	AATTGAAACCAGACGAAATGCA
dd_Smed_v4_6450_0_1	SMESG000034548	mettl3_R	AGGTGTGTGGTTGCAGAGG
dd_Smed_v4_7426_1	SMESG000019097	hakai_F	ACGGTCAATGTCTCGTAGCC
dd_Smed_v4_7426_1	SMESG000019097	hakai_R	GCACGGCATCATGTTTAGG
dd_Smed_v6_2322_4	SMESG000043489	wtap_F	CCCGATGAAATGGTGAAATC
dd_Smed_v6_2322_4	SMESG000043489	wtap_R	TCGGGATCATCCTCTTCATC

dd_Smed_v4_3491_0_1	SMESG000018260	ythdc-1_F	CAGTCATCTCCCAATGTTGACG
dd_Smed_v4_3491_0_1	SMESG000018260	ythdc-1_R	ACAAACCGCAATAATTGTAACCA
dd_Smed_v6_3194_0_1	SMESG000053627	Intestine marker_F	GATCGAGGAAATCTTGACGAAC
dd_Smed_v6_3194_0_1	SMESG000053627	Intestine marker_R	GTGAATACTTCAGGAGCCATCC
dd_Smed_v6_175_0_1	SMESG000049722	*cathepsin*+ marker_F	ACGATCTCGGAAAACATTCG
dd_Smed_v6_175_0_1	SMESG000049722	*cathepsin*+ marker_R	AGCCATCGTAGCTATTCCACA
dd_Smed_v6_701_0_1	SMESG000077782	collagen F	GTGGAGAACTTGGAGCAAGC
dd_Smed_v6_701_0_1	SMESG000077782	collagen R	AACACCAGCATCTCCTGGAC
dd_Smed_v6_4575_0_1	SMESG000068883	Protonephridia_F	CAACCCCAACGCGATACAAT
dd_Smed_v6_4575_0_1	SMESG000068883	Protonephridia_R	CAGAAAGAAAGAGCCTCCGC
dd_Smed_v4_4676_1	SMESG000074445	*kiaa1429* non-overlapping_F	AGGACGTTTCCAGCAATGAG
dd_Smed_v4_4676_1	SMESG000074445	*kiaa1429* non-overlapping_R	CATGGTTCGCCTTGGATTAG
dd_Smed_v6_2331_0_1	SMESG000068192	*CHD4*_F	CGCGAGCATTTCATCTTGTA
dd_Smed_v6_2331_0_1	SMESG000068192	*CHD4*_R	AAACGATGGGCTTCATCAAC
dd_Smed_v6_2065_0_1	NA	*rbAp48*_F	GAATGGCCAAGCTTAACTGC
dd_Smed_v6_2065_0_1	NA	*rbAp48*_R	TTTGGGCTCCAACTGAAATC
dd_Smed_v6_115_0_1	SMESG000066413	Intestine marker_F	CAGTGCTTGCCGTCTGTCTA

dd_Smed_v6_115_0_1	SMESG000066413	Intestine marker_R	TAGCAACCAGTGCATTGAGC
dd_Smed_v4_72_0_1	NA	Intestine marker_F	AATGTTGGGATGCTGCAGTT
dd_Smed_v4_72_0_1	NA	Intestine marker_R	CAAAACGCAGGGCTCATACT
dd_Smed_v4_75_0_1	SMESG000027140	Intestine marker_F	TGCCGTTATGAACATGATTTTCG
dd_Smed_v4_75_0_1	SMESG000027140	Intestine marker_R	ACACAAAATATCGCATCCTGCC
dd_Smed_v4_888_0_1	SMESG000073691	Intestine marker_F	TCTTGGACTTCATCGACTTTCT
dd_Smed_v4_888_0_1	SMESG000073691	Intestine marker_R	AACTGGTTTTCGTTGTCTACAAA
dd_Smed_v6_1837_0_1	SMESG000016779	*kiaa1429*-cluster #1 F	TTGGAACAGACCACTGGTGA
dd_Smed_v6_1837_0_1	SMESG000016779	*kiaa1429*-cluster #1 R	AACGACGACCTTTCCAACTG
dd_Smed_v6_585_0_1	SMESG000062836	*kiaa1429*-cluster #2 F	ACAAGTGCAATGCCCGTAAC
dd_Smed_v6_585_0_1	SMESG000062836	*kiaa1429*-cluster #2 R	CCTCATTTCCCACGGTATCA

103 **Appendix Table S3. Inline m6A-seq2 barcode sequences**

Condition	**Replicate**	**Barcode sequence**	**Experiment**
kiaa1429 (RNAi)	#1	NNNCCAGTCT	m6A-seq2
kiaa1429 (RNAi)	#2	NNNCCTCCGT	m6A-seq2
kiaa1429 (RNAi)	#3	NNNGCACTCT	m6A-seq2
unc22 (RNAi), Control	#1	NNNGCCCATT	m6A-seq2
unc22 (RNAi), Control	#2	NNNGGCCTCT	m6A-seq2
unc22 (RNAi), Control	#3	NNNTATCACT	m6A-seq2

unc22 (RNAi), Control *Ime4*Δ/Δ/*Ndt80*Δ/Δ	#1	NNNGTTTGCT	Species mixing m6A-seq2
unc22 (RNAi), Control *Ime4*Δ/Δ/*Ndt80*Δ/Δ	#2	NNNCCTTAGT	Species mixing m6A-seq2
kiaa1429 (RNAi) *Ndt80*Δ/Δ	#1	NNNCAACTGT	Species mixing m6A-seq2
kiaa1429 (RNAi) *Ndt80*Δ/Δ	#2	NNNCACCTCT	Species mixing m6A-seq2

104 **Appendix Table S4. Mapping of library data to the planarian transcriptome**

Condition	**Type**	**Mapping ratio**	**Unmapped read ratio**
kiaa1429 (RNAi)	Input	0.871783	0.128217
kiaa1429 (RNAi)	pulldown	0.876082	0.123918
kiaa1429 (RNAi)	Input	0.872294	0.127706
kiaa1429 (RNAi)	pulldown	0.88221	0.11779
kiaa1429 (RNAi)	Input	0.88312	0.11688
kiaa1429 (RNAi)	pulldown	0.885548	0.114452
Control	Input	0.886476	0.113524
Control	pulldown	0.895546	0.104454
Control	Input	0.840594	0.159406
Control	pulldown	0.891526	0.108474
Control	Input	0.882362	0.117638

Control	pulldown	0.893802	0.106198

105

Appendix Table S5. qPCR primers used in this study

Gene	Sequence	Contig	Gene model
*11930*_F	GAGGATTCTCCTATTAAACCCAAC	dd_smed_v6_11930_0_1	SMESG000016771
*11930*_R	CGTGCGATCAGCGTTTATATT	dd_smed_v6_11930_0_1	SMESG000016771
*gapdh*_F*	TCTTCCCCAACCAATTTTCTGTTCTG	dd_smed_v6_78_0_1	SMESG000052353
*gapdh*_R*	CCGAATATTTTATTTGGCTCTTCCTCCA	dd_smed_v6_78_0_1	SMESG000052353
*smedwi1*_F*	GTCTCAGAAAACAACTAAAGGTACAGCA	dd_smed_v6_659_0_1	SMESG000036375
*smedwi1*_R*	TGCTGCAATACACTCGGAGACA	dd_smed_v6_659_0_1	SMESG000036375
*mettl3*_F	ACAGTTTTGTCACTATGGTACC	dd_Smed_v6_6450_0_1	SMESG000034548
*mettl3*_R	GAAAGCAAGTATTGAGAAATGAAC	dd_Smed_v6_6450_0_1	SMESG000034548
*wtap*_F	GGAGCTGATTTCCATTAAGAATG	dd_Smed_v6_2322_0_4	SMESG000043489
*wtap*_R	TCGATATTGATTTTGCACACTG	dd_Smed_v6_2322_0_4	SMESG000043489
*3194*_F	CATTTGTGCTATTACGGTGTC	dd_smed_v6_3194_0_1	SMESG000053627
*3194*_R	GGTGAGTTCATTAGTGTACAG	dd_smed_v6_3194_	SMESG000053

	C	0_1	627
*h2b*_F	TGTAAGGTTGATTTTACCAGGAG	dd_smed_v6_4808_0_1	SMESG000067906
*h2b*_R	CCCGTGTACTTAGTAACAGC	dd_smed_v6_4808_0_1	SMESG000067906
*rbm15*_F	GGCGAATTCAATGTCAAAGTAG	dd_Smed_v6_7282_0_1	SMESG000001341
*rbm15*_R	AGCCAAAACAGGATCAACTG	dd_Smed_v6_7282_0_1	SMESG000001341

* Primer sequences previously published (van Wolfswinkel *et al*, 2014).

Supplementary References

Guo T, Peters AHFM & Newmark PA (2006) A *bruno-like* gene is required for stem cell maintenance in planarians. *Dev Cell* 11: 159–169

Marques JG, Gryder BE, Pavlovic B, Chung Y, Ngo QA, Frommelt F, Gstaiger M, Song Y, Benischke K, Laubscher D, *et al* (2020) NuRD subunit CHD4 regulates super-enhancer accessibility in rhabdomyosarcoma and represents a general tumor dependency. *eLife* 9

Wilczewski CM, Hepperla AJ, Shimbo T, Wasson L, Robbe ZL, Davis IJ, Wade PA & Conlon FL (2018) CHD4 and the NuRD complex directly control cardiac sarcomere formation. *Proc Natl Acad Sci USA* 115: 6727–6732

van Wolfswinkel JC, Wagner DE & Reddien PW (2014) Single-cell analysis reveals functionally distinct classes within the planarian stem cell compartment. *Cell Stem Cell* 15: 326–339