Poly(A) tail length and mRNA 3′ terminal uridylation sculpt the mammalian maternal transcriptome

Marcos Morgan[1,2*], Christian Much[1,2*], Monica DiGiacomo[2], Chiara Azzi[2], Ivayla Ivanova[1,2], Dimitrios M. Vitsios[4], Jelena Pistolic[3], Paul Collier[3], Pedro Moreira[1,2], Vladimir Benes[3], Anton J. Enright[4] and Dónal O'Carroll[1,2]

1. MRC Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, 5 Little France Drive, Edinburgh, EH16 4UU, UK.
2. European Molecular Biology Laboratory (EMBL), Mouse Biology Unit, Via Ramarini 32, Monterotondo Scalo, 00015, Italy.
3. European Molecular Biology Laboratory (EMBL), Genomics Core Facility, Meyerhofstraße 1, Heidelberg, 69117, Germany.
4. European Molecular Biology Laboratory (EMBL), European Bioinformatics Institute (EBI), Wellcome Genome Campus, Hinxton, Cambridge, CB10 1SD, UK.

*Equally contributing authors.
Corresponding author:

Dónal O'Carroll
MRC Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, 5 Little France Drive, Edinburgh, EH16 4UU, UK.

Tel +44 131 6519631
Email: donal.ocarroll@ed.ac.uk

Keywords: maternal transcriptome, mRNA uridylation, TUT4, TUT7, oogenesis and germ line.

A fundamental principle in biology is that the program for early development is established during oogenesis in the form of the maternal transcriptome. How the maternal transcriptome acquires the appropriate content and dosage of transcripts is not fully understood. Here we show that TUT4/7-mediated mRNA 3′ terminal uridylation sculpts the mouse maternal transcriptome by eliminating transcripts during oocyte growth. TUT4/7-mediated uridylation is essential for both oocyte maturation and fertility. In comparison to somatic cells, the oocyte transcriptome displays shorter poly(A) tail length and a high relative proportion of terminal oligo-uridylation. TUT4/7 deletion leads to the accumulation of a cohort of transcripts with a high frequency of very short poly(A) tails and their loss of 3′ oligo-uridylation. In contrast, TUT4/7-deficiency does not alter gene expression in a variety of somatic cells. In summary, we show essential and specific functions for poly(A) tail length and 3′ terminal uridylation in sculpting a functional maternal transcriptome.

The early stages of zygotic development are characterized by the lack of transcription and thus gene expression is instructed by the maternally-deposited transcriptome[1]. In mammals, the maternal transcriptome is built during folliculogenesis that encompasses the growth phase of oogenesis and culminates in ovulation[2]. This process initiates when a clutch of primordial oocytes begins to grow coinciding with the expansion of surrounding somatic cells that together will form follicles. The stages of growing oocytes are classified sequentially as primary, secondary, early antral or late antral depending on morphology and increasing cell size (Fig. 1a). The mature oocyte upon ovulation is competent to support fertilization as well as development[3]; this competence is largely defined by the maternal transcriptome[4,5]. The maternal mRNA in growing oocytes is extremely stable[6,7], however, each stage of growth is characterized by a distinct transcriptome[4]. This implies that growing oocytes not only require

the ability to accumulate mRNA but also to mediate its selective and limited degradation. How the growing oocyte eliminates transcripts to accrue the appropriate content and dosage of mRNAs defining a functional maternal transcriptome remains largely unknown.

Poly(A) tail length and 3′ terminal mRNA uridylation are key determinants of mRNA turnover[8–10]. It was recently demonstrated in human cell lines that approximately a fifth of transcripts with very short poly(A) tails (< 25 nucleotides) is 3′ uridylated and that this nucleotide addition optimizes the mRNA for decay[9,11]. The shortening of a poly(A) tail below ~27 nucleotides results in the loss of the stabilizing PABP binding[12,13] and the subsequent recruitment of the terminal uridylyl transferases 4 (TUT4) and TUT7 (TUT4/7) that mediate mRNA 3′ uridylation[9]. LSM (like-Sm) proteins are central regulators of RNA metabolism and have a binding preference for terminal uridylyl residues[14], thus their recruitment sets in motion the 5′-3′ XRN1 decay pathway through the recruitment of DCP1/2 mRNA decapping enzymes[15]. TUT4/7-mediated precursor-miRNA 3′ uridylation functions in the regulation of miRNA biogenesis[16–22]. Here, oligo-uridylation targets pre-let-7 for degradation through the recruitment of DIS3L2[23,24], a 3′-5′ ribonuclease that specifically recognizes RNA with an oligo(U) tail[25]. Terminal uridylation may also optimize mRNA for 3′-5′ exosome and/or DIS3L2-mediated decay[9,26,27]. The contribution of TUT4/7-mediated 3′ uridylation to mRNA degradation *in vivo* and physiology is undetermined. A transcriptome-wide analysis of poly(A) tail length in growing oocytes is lacking, however, analysis of several individual messages in germinal vesicle (GV) (late antral/preovulatory) oocytes indicates an overall short poly(A) tail length of ~50 nucleotides[28–31]. Here we sought to understand if poly(A) tail length and mRNA 3′ uridylation function in the formation of the maternal transcriptome.

The GV oocyte transcriptome is characterized by a high relative frequency of oligo-uridylation

Firstly, to explore a possible contribution of TUT4/7-mediated mRNA uridylation during oocyte growth, we generated epitope-tagged alleles for both genes (Extended Data Fig. 1). Both TUT4 and TUT7 proteins were expressed from the primordial through to late antral stages of oogenesis (Fig. 1b). We next sought to understand the terminal uridylation profile and poly(A) tail features (3′-terminome) of the GV oocyte transcriptome. We therefore performed TAIL-seq[11] with a slight modification to correct for poly(A) tail length recovery bias (Methods and Extended Data Fig. 2) from GV oocytes as well as from somatic mouse tissues/cells (liver, bone marrow, mouse embryonic fibroblasts (MEFs) and embryonic stem cells (ESCs)). TAIL-seq enables the simultaneous determination of poly(A) tail length and the 3′ terminal nucleotides[11]. Oocytes had the poly(A) tail length distribution with the shortest mode length (~65 nucleotides) (Fig. 1c). While the overall amount of terminal uridylation in the respective transcriptomes differed (Fig. 1d), a distinguishing feature of the GV oocyte was that it presented the highest relative ratio of oligo-to-mono-uridylation among the respective transcriptomes (Fig. 1e, f and Extended Data Table 1).

TUT4/7 and specifically their uridylation activity are intrinsically required for oocyte maturation

Intrigued by the unique features of the GV oocyte terminome, we sought to use conditional mutagenesis to explore the function of TUT4/7-mediated uridylation during oogenesis. The fact that TUT4-deficient mice manifest growth defects[32] but are fertile indicates possible redundancy with TUT7. We therefore engineered conditional alleles for both genes in the mouse (Extended Data Fig. 3) and combined these *Tut4*^{Fl} and *Tut7*^{Fl} alleles with a Zp3Cre transgene that deletes at the secondary oocyte stage[33] to generate control *Tut4*$^{+/Fl}$; *Tut7*$^{+/Fl}$;

Zp3Cre Tg^{+} *or Tut4*$^{+/+}$*; Tut7*$^{+/+}$*; Zp3Cre Tg*$^{+}$ (*Tut4/7*^{CTL}) and experimental *Tut4*$^{Fl/Fl}$*; Tut7*$^{Fl/Fl}$*; Zp3Cre Tg*$^{+}$ (*Tut4/7*^{cKO}) mice. *Tut4/7*^{cKO} mice were infertile, although they ovulated normal numbers of oocytes and corpora lutea were observed (Fig. 2a-c). The morphology of antral follicles in the *Tut4/7*^{cKO} mice was also normal with a slight decrease in the frequency of surrounded nucleolus state oocytes observed (Extended Data Fig. 4). To understand why the mutant oocytes were defective, we hormone primed and set up *Tut4/7*^{CTL} and *Tut4/7*^{cKO} females with wild type males and placed isolated oocytes/zygotes in culture to monitor pre-implantation development *in vitro*. Over the course of five days, the bulk of *Tut4/7*^{CTL} embryos developed until the blastocyst stage; although sperm contact was observed with *Tut4/7*^{cKO} oocytes, the majority of them did not develop further (Fig. 2d). A small fraction of the *Tut4/7*^{cKO} fertilized oocytes were observed at the 2-pronuclei stage but never developed past 2-cell embryos (Fig. 2d). In summary, *Tut4/7*^{cKO} oocytes are incompetent to support early embryonic development. This phenotype could arise from failure to complete meiosis I properly thus generating aberrant MII oocytes, or alternatively, *Tut4/7*^{cKO} oocytes complete meiosis I correctly but the resulting MII oocytes are not capable of supporting fertilization. To discriminate between these two possibilities, *Tut4/7*^{CTL} and *Tut4/7*^{cKO} females were hormone primed and ovulated oocytes were isolated. The majority of *Tut4/7*^{cKO} oocytes failed to complete meiosis I properly and multiple phenotypic abnormalities were observed, including abnormal spindle morphology with microtubule asters (Fig. 2e-II), telophase arrest (Fig. 2e-III) as well as aberrant MII oocytes with chromosomes dispersed along the spindle (Fig. 2e-IV). This defective meiosis could also be observed with *in vitro* cultured GV oocytes with the majority of oocytes failing to extrude the first polar body (Fig. 2f, g). While the above data identify a critical function for TUT4/7 in oocyte maturation, we wanted to formally understand if the uridylation activity of TUT4/7 is the principal function by which these large multidomain enzymes support meiosis I. We therefore took advantage of a

previously characterized mutation in the catalytic triad (DDD to AAD) that renders TUT4 enzymatically inactive[21,34] and engineered a *Tut4*^{AAD} allele in the mouse (Extended Data Fig. 5). To understand if the defective meiosis I phenotype is dependent upon the uridylation activity of TUT4, we combined the *Tut4*^{Fl}, *Tut7*^{Fl}, *Tut4*^{AAD} and *Zp3Cre* alleles to generate experimental *Tut4*$^{Fl/AAD}$; *Tut7*$^{Fl/Fl}$; *Zp3Cre Tg*$^{+}$ (*Tut4/7*^{cAAD}) mice, where only a single copy of the catalytic-dead TUT4^{AAD} protein is expressed during oogenesis. While a single copy of TUT4 supported oogenesis, the expression of the catalytic-dead TUT4^{AAD} protein phenocopied the conditional ablation of TUT4/7 (Fig. 2a-g). In summary, the function of TUT4/7 and specifically their uridylation activity are intrinsically required during oocyte growth to complete meiosis I and generate functional MII oocytes.

TUT4/7 are required for transcript elimination

To understand if a defective maternal transcriptome underpinned the *Tut4/7*^{cKO} phenotype, we analyzed gene expression in *Tut4/7*^{CTL} and *Tut4/7*^{cKO} GV oocytes, which represent the terminal stages of oocyte growth just prior to the onset of meiotic maturation and ovulation. The analysis of respective transcriptomes revealed the deregulation of many genes (Fig. 3a). The majority of deregulated genes were upregulated in the absence of TUT4/7 (Fig. 3a and Extended Data Fig. 6), an expected outcome of removing a component of an RNA degrading pathway. The same deregulation of gene expression was observed in *Tut4/7*^{cAAD} as in *Tut4/7*^{cKO} GV oocytes (Extended Data Fig. 7a and b). Gene ontology analysis of the deregulated genes did not identify any germline-specific processes but rather revealed generic, mostly metabolic, pathways (Extended Data Fig. 7c). We therefore sought to probe the relationship between the deregulated genes and oogenesis. To this end, we analyzed five growing oocyte transcriptomes: primordial, primary, secondary, early and late antral oocytes[4]. We applied the Markov clustering (MCL) algorithm[35] to group genes in terms of

expression patterns across oocyte growth (Fig. 3b). We next addressed if we could detect enrichment or depletion of the TUT4/7-dependent upregulated genes in any of the gene expression patterns. We identified significant enrichment in clusters where genes are downregulated across oogenesis (Fig. 3c). Conversely, depletion of the gene set was observed in clusters where genes were upregulated or remained unchanged across development (Fig. 3c). Thus, TUT4/7 and specifically their uridylation activity are required to degrade transcripts during oocyte growth; in essence they sculpt a functional maternal transcriptome.

TUT4/7-mediated oligo-uridylation targets transcripts with short poly(A) tails for degradation

Next we sought to understand why a group of defined transcripts is dependent upon TUT4/7-mediated uridylation during oocyte growth. To this end, we performed TAIL-seq on *Tut4/7*^{cKO} GV oocytes and analyzed the 3′-terminome of the TUT4/7-dependent upregulated genes and unchanged genes. Interestingly, upon deletion of TUT4/7, the fraction of the TUT4/7-dependent upregulated genes with short poly(A) tails dramatically increased (Fig. 4a), indicating that the failure to degrade the transcripts resulted in their accumulation with short poly(A) tails. TUT4/7-deficiency had little impact on terminal mono-uridylation in transcripts with short poly(A) tails (Fig. 4b and Extended Data Table 2). The oligo-uridylation of the TUT4/7-dependent upregulated genes with short poly(A) tails was dependent upon TUT4/7 with not a single oligo-uridylated read for TUT4/7-dependent upregulated genes observed in *Tut4/7*^{cKO} GV oocyte libraries (Fig. 4c and Extended Data Table 2). The ratio of oligo- to mono-uridylation of transcripts with short poly(A) tails, a defining feature of the GV oocyte transcriptome, was dependent upon TUT4/7 function (Fig. 4c). Thus, the loss of TUT4/7-dependent oligo-uridylation in GV oocytes results in the accumulation of transcripts with short poly(A) tails.

A selective function for TUT4/7-mediated uridylation in sculpting the maternal transcriptome

One can argue that the loss of a central RNA degradation targeting pathway will inevitably be detrimental to oogenesis or indeed any cell type or tissue. We therefore sought to understand if TUT4/7-mediated uridylation is a ubiquitous regulator of mRNA degradation. To this end, we combined the ubiquitously-expressed tamoxifen inducible Cre ($R26^{ERTCre}$) allele with the $Tut4^{Fl}$ and $Tut7^{Fl}$ alleles to generate control $Tut4^{+/Fl}$; $Tut7^{+/Fl}$; $R26^{ERTCre/+}$ and experimental $Tut4^{Fl/Fl}$; $Tut7^{Fl/Fl}$; $R26^{ERTCre/+}$ mice. ESCs and MEFs were derived from these mice and tamoxifen treatment used to induce deletion, generating control $Tut4^{+/-}$; $Tut7^{+/-}$; $R26^{ERTCre/+}$ ($Tut4/7^{CTL}$) and experimental induced $Tut4^{-/-}$; $Tut7^{-/-}$; $R26^{ERTCre/+}$ ($Tut4/7^{iKO}$) cell lines (Extended Data Fig. 8 and 9). The loss of TUT4/7 did not impact ESC/MEF proliferation, ESC pluripotency or differentiation (Extended Data Fig. 8 and 9). We also administered tamoxifen *in vivo* to generate $Tut4/7^{CTL}$ and $Tut4/7^{iKO}$ adult mice; sampling of liver and bone marrow confirmed complete deletion in the respective tissues (Extended Data Fig. 10), and these mice appeared healthy up to several months after deletion. We next performed TAIL-seq on the respective transcriptomes to gain insight into terminal mRNA uridylation *in vivo*. The loss of TUT4/7 did not grossly impact the poly(A) tail length profile of the transcriptomes in the respective cell types (Fig. 5a). However, the loss of TUT4/7 reduced but did not eliminate terminal mRNA uridylation (Fig. 5b). The degree of this reduction ranged from ~2 to 5-fold examining all transcripts independent of poly(A) length. As observed from the analysis of immortalized cell lines[9], the fold reduction in terminal uridylation was greater when short poly(A) tails were selectively analyzed, however, the fold reduction again ranged from the 2.5 to 7.5-fold, with the exception of bone marrow that displayed a 28-fold reduction (Fig. 5b). Both mono-uridylation and to a higher extent oligo-uridylation were reduced in all of the respective transcriptomes (Extended Data Fig. 11).

Furthermore, the reduction of uridylation in the respective transcriptomes did not significantly affect other terminal mRNA modifications (Extended Data Fig. 12). In contrast to oocytes, the miRNA pathway is functional in somatic cells[36,37]. The loss of TUT4/7 resulted in the reduction of terminal uridylation with a more severe reduction observed for let-7 family members in all somatic cells analyzed (Extended Data Fig. 13a). No significant impact on terminal miRNA adenylation, guanylation or cytidylation was observed in the respective *Tut4/7*^{iKO} cells/tissues (Extended Data Fig. 13b-d). TUT4/7-deficiency overall had a very modest impact on miRNA expression levels (Fig. 5c and Extended Data Fig. 14). In ESCs, where LIN28a is expressed and TUT4/7-mediated oligo-uridylation inhibits pre-let-7 processing[19,21], a modest increase in let-7 dosage was observed in the absence of TUT4/7. The opposite effect was seen in liver, bone marrow and MEFs (Fig. 5c and Extended Data Fig. 14), where LIN28 is not expressed and TUT4/7-mediated mono-uridylation promotes specific pre-let7 processing[17]. Finally, the loss of TUT4/7 with the consequent reduction of terminal mRNA uridylation and modest alterations to the miRNome did not have any appreciable impact on the gene expression within any of the cell types/tissues analyzed (Fig. 5d and Extended Data Fig. 15). In summary, TUT4/7-mediated uridylation is not an essential requirement for mRNA degradation in somatic cells.

Discussion

Here we describe mRNA poly(A) tail length and 3′ terminal nucleotide additions in a variety of somatic primary mouse cell lines and tissues as well as oocytes. We observe that each tissue/cell line has a defined signature in terms of poly(A) tail length profile as well as the frequency and preference of terminal nucleotide additions. Furthermore, the reduction in terminal uridylation in TUT4/7-deficient cells does not alter terminal cytidylation or guanylation, thus the addition of these nucleotides is likely independent of uridylation. While

TUT4/7-mediated uridylation is essential for the regulation of gene expression during oogenesis and in HeLa cells[9], it has no appreciable impact on gene expression in any of the somatic cell/tissue types analyzed. This could be due to the fact that loss of TUT4/7 reduces but does not abolish terminal uridylation in the somatic cells. This fact indicates that other TUTases can uridylate mRNA, possibly providing redundancy within the system. Alternatively, the analyzed somatic cell types may not be reliant on 3′ uridylation for RNA degradation under steady state conditions. Oocytes and HeLa cells share a potentially important commonality as they both have a relatively high frequency of 3′ mRNA oligo-uridylation[9,11]; as discussed below, oligo-uridylation could be the potent degradation signal[9]. However, given that TUT4/7 are broadly expressed, one cannot exclude a function for TUT4/7-mediated mRNA 3′ uridylation beyond steady state in managing somatic cell stress responses[27].

A widely recognized feature in building a maternal transcriptome is the necessity to stabilize RNA that is achieved by stabilizing RNA binding proteins such as MSY2[38] and the downregulation of RNA degradation pathways[31,39]. However, limited degradation is also required to define the maternal transcriptome inferred by the distinct transcriptomes of growing oocyte stages[4] and data presented herein. The consequence of failing to degrade transcripts during oocyte growth results in the overexpression of ~750 genes and generates a dysfunctional maternal transcriptome, poisoning oocyte maturation. A very short poly(A) tail likely provides the recognition signal to engage the TUT4/7-mediated 3′ uridylation pathway. Indeed, a poly(A) tail length of <27 nucleotides leads to a loss of PABP occupancy and is a proven substrate for TUT4/7 binding and subsequent uridylation[9]. In addition, the failure to degrade the TUT4/7-dependent transcripts results in their accumulation with ~10 nucleotides poly(A) tails. The derailing of meiotic maturation in *Tut4/7*^{cKO} mice may arise from alterations to the proteome and/or the accumulation of these transcripts with short poly(A)

tails. Oligo-uridylation appears to be a primary signal that instructs transcript degradation in oocytes, as evidenced by the observation that TUT4/7-deficiency modestly affects terminal mono-uridylation but abolishes oligo-uridylation in the TUT4/7-targeted transcripts with short poly(A) tails. An increase in poly(A) length of specific maternal transcripts is associated with their translation during mouse oocyte maturation or early zygotic development[28–30,40]. Indeed, this positive correlation between poly(A) tail length and translation efficiency is a conserved feature of animal development[41–43]. The function of mRNA uridylation in gametogenesis and embryonic development is less well understood, although there is little terminal uridylation found in the *Drosophila* early embryo[41]. Here we show a function for 3′ terminal uridylation and poly(A) tail length in the formation of the maternal mammalian transcriptome. Thus, poly(A) tail length regulation is important for both defining and activating the maternal transcriptome.

METHODS SUMMARY

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

1. Tadros, W. & Lipshitz, H. D. The maternal-to-zygotic transition: a play in two acts. *Development* **136,** 3033–3042 (2009).
2. Svoboda, P., Franke, V. & Schultz, R. M. *Sculpting the Transcriptome During the Oocyte-to-Embryo Transition in Mouse*. *Current Topics in Developmental Biology* **113,** (Elsevier Inc., 2015).
3. Eppig, J. J. & Schroeder, a C. Capacity of mouse oocytes from preantral follicles to undergo embryogenesis and development to live young after growth, maturation, and

fertilization in vitro. *Biol. Reprod.* **41,** 268–276 (1989).

4. Pan, H., O'brien, M. J., Wigglesworth, K., Eppig, J. J. & Schultz, R. M. Transcript profiling during mouse oocyte development and the effect of gonadotropin priming and development in vitro. *Dev. Biol.* **286,** 493–506 (2005).
5. Ma, J.-Y. *et al.* Maternal factors required for oocyte developmental competence in mice: Transcriptome analysis of non-surrounded nucleolus (NSN) and surrounded nucleolus (SN) oocytes. *Cell Cycle* **12,** 1928–1938 (2013).
6. Brower, P. T., Gizang, E., Boreen, S. M. & Schultz, R. M. Biochemical studies of mammalian oogenesis: Synthesis and stability of various classes of RNA during growth of the mouse oocyte in vitro. *Dev. Biol.* **86,** 373–383 (1981).
7. De Leon, V., Johnson, A. & Bachvarova, R. Half-lives and relative amounts of stored and polysomal ribosomes and poly(A)+ RNA in mouse oocytes. *Dev. Biol.* **98,** 400–408 (1983).
8. Rissland, O. S. & Norbury, C. J. Decapping is preceded by 3' uridylation in a novel pathway of bulk mRNA turnover. *Nat. Struct. Mol. Biol.* **16,** 616–23 (2009).
9. Lim, J. *et al.* Uridylation by TUT4 and TUT7 Marks mRNA for Degradation. *Cell* **159,** 1365–1376 (2014).
10. Mullen, T. E. & Marzluff, W. F. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5′ to 3′ and 3′ to 5′. *Genes Dev.* **22,** 50–65 (2008).
11. Chang, H., Lim, J., Ha, M. & Kim, V. N. TAIL-seq: Genome-wide Determination of Poly(A) Tail Length and 3' End Modifications. *Mol. Cell* **53,** 1044–1052 (2014).
12. Baer, B. W. & Kornberg, R. D. The protein responsible for the repeating structure of cytoplasmic poly(A)-ribonucleoprotein. *J. Cell Biol.* **96,** 717–721 (1983).
13. Eliseeva, I. A., Lyabin, D. N. & Ovchinnikov, L. P. Poly (A) Binding Proteins :

Structure , Domain Organization , and Activity Regulation. **78,** (2013).

14. Chowdhury, A., Mukhopadhyay, J. & Tharun, S. The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. *RNA* **13,** 998–1016 (2007).

15. Song, M.-G. & Kiledjian, M. 3' Terminal oligo U-tract-mediated stimulation of decapping. *RNA* **13,** 2356–65 (2007).

16. Kim, B. *et al.* TUT7 controls the fate of precursor microRNAs by using three different uridylation mechanisms. *EMBO J.* **34,** 1801–1815 (2015).

17. Heo, I. *et al.* Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. *Cell* **151,** 521–532 (2012).

18. Lehrbach, N. J. *et al.* LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in Caenorhabditis elegans. *Nat. Struct. Mol. Biol.* **16,** 1016–1020 (2009).

19. Heo, I. *et al. Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA*. *Molecular Cell* **32,** (2008).

20. Heo, I. *et al.* TUT4 in Concert with Lin28 Suppresses MicroRNA Biogenesis through Pre-MicroRNA Uridylation. *Cell* **138,** 696–708 (2009).

21. Hagan, J. P., Piskounova, E. & Gregory, R. I. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nat. Struct. Mol. Biol.* **16,** 1021–1025 (2009).

22. Thornton, J. E., Chang, H.-M., Piskounova, E. & Gregory, R. I. Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). *Rna* **18,** 1875–1885 (2012).

23. Chang, H.-M., Triboulet, R., Thornton, J. E. & Gregory, R. I. A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. *Nature* **497,** 244–8 (2013).

24. Ustianenko, D. *et al.* Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. *RNA* **19,** 1632–8 (2013).

25. Malecki, M. *et al.* The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. *EMBO J.* **32,** 1842–54 (2013).

26. Lubas, M. *et al.* Exonuclease hDIS3L2 specifies an exosome-independent 3'-5' degradation pathway of human cytoplasmic mRNA. *EMBO J.* **32,** 1855–68 (2013).

27. Thomas, M. P. *et al.* Apoptosis Triggers Specific, Rapid, and Global mRNA Decay with 3′ Uridylated Intermediates Degraded by DIS3L2. *Cell Rep.* 1–11 (2015). doi:10.1016/j.celrep.2015.04.026

28. Huarte, J., Belin, D., Vassalli, A., Strickland, S. & Vassalli, J. D. Meiotic maturation of mouse oocytes triggers the translation and polyadenylation of dormant tissue-type plasminogen activator mRNA. *Genes Dev.* **1,** 1201–1211 (1987).

29. Tay, J., Hodgman, R. & Richter, J. D. The Control of Cyclin B1 mRNA Translation during Mouse Oocyte Maturation. doi:10.1006/dbio.2000.9669

30. Chen, J. *et al.* Genome-wide analysis of translation reveals a critical role for deleted in azoospermia-like (Dazl) at the oocyte-to-zygote transition. *Genes Dev.* **25,** 755–66 (2011).

31. Ma, J., Fukuda, Y. & Schultz, R. M. Mobilization of Dormant Cnot7 mRNA Promotes Deadenylation of Maternal Transcripts During Mouse Oocyte Maturation. *Biol. Reprod.* **93,** 48–48 (2015).

32. Jones, M. R. *et al.* Zcchc11 uridylates mature miRNAs to enhance neonatal IGF-1 expression, growth, and survival. *PLoS Genet.* **8,** e1003105 (2012).

33. de Vries, W. N. *et al.* Expression of Cre recombinase in mouse oocytes: a means to study maternal effect genes. *Genesis* **26,** 110–2 (2000).

34. Jones, M. R. *et al.* Zcchc11-dependent uridylation of microRNA directs cytokine

expression. *Nat. Cell Biol.* **11,** 1157–1163 (2009).

35. Freeman, T. C. *et al.* Construction, visualisation, and clustering of transcription networks from microarray expression data. *PLoS Comput. Biol.* **3,** 2032–2042 (2007).

36. Suh, N. *et al.* Report MicroRNA Function Is Globally Suppressed in Mouse Oocytes and Early Embryos. *Curr. Biol.* **20,** 271–277 (2010).

37. Ma, J. *et al.* Report MicroRNA Activity Is Suppressed in Mouse Oocytes. *Curr. Biol.* **20,** 265–270 (2010).

38. Medvedev, S., Yang, J., Hecht, N. B. & Schultz, R. M. CDC2A (CDK1)-mediated phosphorylation of MSY2 triggers maternal mRNA degradation during mouse oocyte maturation. *Dev. Biol.* **321,** 205–215 (2008).

39. Ma, J., Flemr, M., Strnad, H., Svoboda, P. & Schultz, R. M. Maternally Recruited DCP1A and DCP2 Contribute to Messenger RNA Degradation During Oocyte Maturation and Genome Activation in Mouse. *Biol. Reprod.* **88,** 11–11 (2013).

40. Gebauer, F., Xu, W., Cooper, G. M. & Richter, J. D. Translational control by cytoplasmic polyadenylation of c-mos mRNA is necessary for oocyte maturation in the mouse. *EMBO J.* **13,** 5712–20 (1994).

41. Subtelny, A. O., Eichhorn, S. W., Chen, G. R., Sive, H. & Bartel, D. P. Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature* **508,** 66–71 (2014).

42. Lim, J., Lee, M., Son, A., Chang, H. & Kim, V. N. mTAIL-seq reveals dynamic poly (A) tail regulation in oocyte-to-embryo development. 1–12 (2016). doi:10.1101/gad.284802.116.3

43. Eichhorn, S. W. *et al.* mRNA Poly(A)-tail Changes Specified by Deadenylation Broadly Reshape Translation in Drosophila Oocytes and Early Embryos. *Elife* **5,** (2016).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements

The research leading to these results has received funding from the Wellcome Trust (Award 106144). This study was technically supported by EMBL Monterotondo's genome engineering and microscopy core facilities.

Author Contributions M.M and C.M. contributed to the design, execution and analysis of most experiments. M.M. performed most of the molecular biology for library generation and much of the bioinformatics analyses. C.M. performed all inducible conditional genetics in somatic cells/tissues. M.D.G, C.A. and I.I. performed oocyte analysis and collection. D.V. performed the small RNA sequencing analysis. P.M. performed the *in vitro* embryo development experiments. J.P. and V.B. optimized gene expression analysis from oocytes. J.P., P.C. and V.B. optimized spike-ins and sequencing of TAIL-seq libraries. A.J.E. constructed the TAIL-seq pipeline to process the raw sequencing data and oversaw all bioinformatic analyses performed. D.O'C. conceived and supervised this study and wrote the final version of the manuscript.

Author Information All mRNA and miRNA expression data are deposited in ArrayExpress E-MTAB-5056 and E-MTAB-5677, respectively. The primary data and output of the TAIL-seq pipeline can be retrieved from: http://www.ebi.ac.uk/research/enright/morgan-et-al-2017. The scripts used in the analysis of the TAIL-seq output can be found at: https://github.com/marcosmorgan/TAIL-seq. The *Tut4* and *Tut7* alleles will be freely

available on a non-collaborative basis. Reprints and permissions information are available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.O'C. (donal.ocarroll@ed.ac.uk).

Figure Legends

Figure 1 | The oocyte transcriptome is characterized by short poly(A) tails and oligo-uridylation. a, Schematic representation of folliculogenesis and ovulation. **b**, Confocal immunofluorescence micrographs of *Tut4*$^{HA\text{-}GFP/HA\text{-}GFP}$ and *Tut7*$^{HA\text{-}GFP/HA\text{-}GFP}$ ovary sections stained with anti-HA antibody (green) are shown as indicated. Broken white circles highlight the oocyte within the follicle. Insets are magnifications of the oocytes. Scale bars are 30 µm. **c**, Frequency of transcripts plotted against poly(A) tail length in liver, bone marrow (BM), mouse embryonic fibroblasts (MEFs), embryonic stem cells (ESCs) and germinal vesicle (GV) oocyte transcriptomes is shown. Transcripts with poly(A) tail length shorter than 80 nucleotides are plotted. Data points and horizontal lines represent biological replicates' mean and range, respectively. The number of transcripts, genes and percentage of reads with poly(A) tail length shorter than 80 nucleotides are as follows: Liver, 724,498 transcripts, 8,395 genes, 55%; BM, 52,809 transcripts, 5,292 genes, 11%; MEFs, 130,313 transcripts, 6,884 genes, 24%; ESCs, 123,459 transcripts, 7,681 genes, 26%; GV, 1,035,640 transcripts, 11,747 genes, 43%. **d**, Percentage of mRNA 3′ terminal transcriptome uridylation is shown. The number of transcripts uridylated and in parentheses the number of genes represented by those transcripts are shown. **e**, Relative frequency (percentage of total uridylation) of indicated U-tail length for transcripts with short (≤ 30 nucleotides) poly(A) tails from the

indicated tissues/cells is shown. **f**, Ratio of oligo- to mono-uridylation for transcripts with short (≤ 30 nucleotides) poly(A) tails in the indicated tissues/cells is shown. The total numbers of transcripts oligo- and mono-uridylated are shown on the upper and bottom panels, respectively. In parentheses the numbers of genes represented by those transcripts are indicated. Data points represent biological replicates; the bars' heights and vertical lines indicate the mean and range, respectively, for (**d**-**f**).

Figure 2 | Infertility and defective oocyte maturation in *Tut4/7*^{cKO} mice.

a, The number of pups born per plug from *Tut4/7*^{CTL}, *Tut4/7*^{cKO} and *Tut4/7*^{cAAD} mice is shown. The number of animals tested, the mean and s.d. are indicated. (t-test; ***, $p < 0.001$) **b**, The number of ovulated oocytes harvested after hormonal stimulation is shown for mice of the indicated genotypes. Center values and error bars indicate the mean and s.d. respectively. The number (n) of mice analyzed is n=8 for *Tut4/7*^{CTL}; n=7 for *Tut4/7*^{cKO} and n=6 *Tut4/7*^{cAAD} mice. **c**, Representative ovary sections stained with H&E for *Tut4/7*^{CTL}, *Tut4/7*^{cKO} and *Tut4/7*^{cAAD} mice is shown. Corpora lutea are indicated with an asterisk, scale bars indicate 200 µm. **d**, *In vitro* development of oocytes isolated from *Tut4/7*^{CTL}, *Tut4/7*^{cKO} and *Tut4/7*^{cAAD} female mice that had been mated with wild type males of proven fertility. The fraction of oocytes with sperm contact or zygotes at the 2-pronuclei (2PN), 2-cell, 4-cell, 8-cell, morula and blastocyst stages is shown over five consecutive days. The number (n) of mice and oocytes/embryos analyzed is: *Tut4/7*^{CTL}, n=137 embryos, n=5 mice; *Tut4/7*^{cKO}, n=119 embryos, n=6 mice and *Tut4/7*^{cAAD}, n=15 embryos, n=2 mice. **e**, Representative confocal immunofluorescence micrographs showing *Tut4/7*^{CTL} (panel I) and abnormal *Tut4/7*^{cKO} (panel II-IV) MII oocytes stained with anti-tubulin antibody (green) and DNA stained with Hoechst 33342 (red) are presented (left). Scale bars indicate 10 µm. The frequency of abnormal MII oocytes is presented for the respective genotypes (right). **f**, *In vitro* oocyte maturation of

Tut4/7^{CTL}, *Tut4/7*^{cKO} and *Tut4/7*^{cAAD} oocytes. The frequency of germinal vesicle (GV) oocytes, germinal vesicle breakdown (GVBD) oocytes and oocytes with polar bodies (PB) is shown at different time points after collection and culture. The number (n) of oocytes analyzed is: *Tut4/7*^{CTL} n=160 oocytes, n=4 mice; *Tut4/7*^{cKO} n=194 oocytes, n=3 mice and *Tut4/7*^{cAAD} n=183 oocytes, n=3 mice. **g**, Representative confocal immunofluorescence micrographs of *in vitro* maturating *Tut4/7*^{CTL}, *Tut4/7*^{cKO} and *Tut4/7*^{cAAD} oocytes stained with anti-tubulin antibody (green) and DNA stained with Hoechst 33342 (red) and the indicated time points are presented. Scale bars indicate 20 μm.

Figure 3 | TUT4/7 are required for transcript degradation during oocyte growth.

a, Expression scatterplot showing relative average expression of transcripts between *Tut4/7*^{CTL} and *Tut4/7*^{cKO} germinal vesicle (GV) oocytes. Significantly deregulated ($p < 0.01$) genes with a fold change greater than 2 are highlighted in red. **b**, Diagram of oocyte growth stages, with abbreviations used, that have been analyzed (left). The top 30 clusters resulting from the clustering analysis according to gene expression patterns across oocyte growth stages (x-axis) are shown (right). The average expression change for genes within the cluster is indicated by a black line with the standard deviation range indicated in gray. **c**, Enrichment and depletion analysis of *Tut4/7*^{cKO} upregulated transcripts in gene expression clusters (defined in (**b**)); clusters that show significance are presented. The p-values for enrichment are indicated on top blue bar and those for depletion on bottom red bar. The hyper-geometric test was used to establish the significance. n represents the number of transcripts defining the cluster and x the number of *Tut4/7*^{cKO} upregulated transcripts within the cluster.

Figure 4 | Short poly(A) tail length and TUT4/7-mediated oligo-uridylation target transcripts for degradation.

a, Frequency of transcripts plotted against poly(A) tail length in *Tut4/7*^{CTL} and *Tut4/7*^{cKO} germinal vesicle (GV) oocyte transcriptomes is shown. Transcripts upregulated and not upregulated in *Tut4/7*^{cKO} oocytes are indicated in red and black, respectively. Transcripts with poly(A) tail shorter than 80 nucleotides are plotted. Data points and horizontal lines represent biological replicates' mean and range, respectively. The number of transcripts, genes and percentage of reads with poly(A) tail length shorter than 80 nucleotides are as follows: *Tut4/7*^{CTL} not upregulated, 1,046,357 transcripts, 11,177 genes, 44%; *Tut4/7*^{CTL} upregulated, 15,339 transcripts, 599 genes, 46%; *Tut4/7*^{cKO} not upregulated, 175,882 transcripts, 8,816 genes, 39%; *Tut4/7*^{cKO} upregulated, 6,281 transcripts, 554 genes, 59%. **b**, **c**, Quantification of mRNA 3′ terminal mono- (**b**) and oligo-uridylation (**c**) of transcripts upregulated and not upregulated in *Tut4/7*^{cKO} oocytes with short poly(A) tails (≤ 30 nucleotides) from *Tut4/7*^{CTL} and *Tut4/7*^{cKO} GV oocytes are presented. The numbers of transcripts mono- or oligo-uridylated, and in parentheses the number of genes represented by these transcripts, are shown. The bars' heights and vertical lines indicate the mean and range, respectively. **d**, Ratio of oligo- to mono-uridylation for transcripts with short (≤ 30 nucleotides) poly(A) tails from *Tut4/7*^{CTL} and *Tut4/7*^{cKO} GV oocytes is shown. The fold change and the significance are indicated (t-test; ns, $p \geq 0.05$; *, $p < 0.05$) (b-d). The data for the *Tut4/7*^{CTL} are derived from the libraries presented in Fig. 1.

Figure 5 | TUT4/7-dependent uridylation of transcripts in different tissues and cell lines does not impact gene expression.

a, Frequency of transcripts plotted against poly(A) tail length from *Tut4/7*^{CTL} and *Tut4/7*^{iKO} liver, bone marrow (BM), mouse embryonic fibroblasts (MEFs), embryonic stem cells (ESCs) transcriptomes are shown. Transcripts with poly(A) tail shorter than 80 nucleotides are plotted. Data points and horizontal lines represent biological replicates' mean and range,

respectively. The data for the *Tut4/7*^{CTL} are derived from the libraries presented in Fig. 1. For the *Tut4/7*^{iKO} libraries the number of transcripts, genes and percentage of reads with poly(A) tail length shorter than 80 nucleotides are as follows: Liver, 1,604,023 transcripts, 9,986 genes, 55%; BM, 62,284 transcripts, 5,638 genes, 11%; MEFs, 115,624 transcripts, 6,754 genes, 34%; ESCs, 113,808 transcripts, 6,897 genes, 32%. **b,** Quantification of mRNA 3′ terminal uridylation of all transcripts (left panel) and those with short (≤ 30 nucleotides) (center panel) and long (> 30 nucleotides) (right panel) poly(A) tails from *Tut4/7*^{CTL} and *Tut4/7*^{iKO} tissues and cell lines are indicated. The numbers of transcripts uridylated, and in parentheses the number of genes represented by these transcripts, are shown. The bars' heights and vertical lines indicate the mean and range, respectively. The fold reduction in uridylation and the significance are indicated (t-test; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). The data for the *Tut4/7*^{CTL} are derived from the libraries presented in Fig. 1. **c**, miRNA expression scatterplots showing relative average miRNA expression levels between *Tut4/7*^{CTL} and *Tut4/7*^{iKO} tissues and cell lines as indicated. Let-7 family members and other miRNAs with significant ($p < 0.05$) and more than 2-fold change in expression levels are highlighted in red and blue, respectively. The linear regression fit is shown in red. **d**, Expression scatterplots showing relative average expression of transcripts between *Tut4/7*^{CTL} and *Tut4/7*^{iKO} tissues and cell lines as indicated. Significantly deregulated ($p < 0.01$) genes with a fold change greater than 2 are shown in red.

METHODS

Mice and alleles used in this study

For the *Tut4*$^{HA\text{-}GFP}$ allele, we inserted the sequence encoding FLAG-HA$_2$-PreScission-His6x-eGFP (FHpHG) after the endogenous ATG initiation codon within exon 3, the first coding exon of *Tut4*. A targeting construct was recombineered that contained homology arms and a

loxP flanked neomycin cassette 3′ of exon 3. Southern blotting of genomic NheI-digested DNA from individual ESC-derived clones with a 3′ probe was used to identify homologous recombinants. A 12.0-kb DNA fragment corresponds to the wild type *Tut4* locus; integration of the loxP flanked neomycin cassette 3′ of exon 3 introduced an additional NheI site, thus decreasing the size of the NheI DNA fragment to 7.1 kb in the targeted allele. Cre-mediated recombination and excision of the loxP flanked neomycin cassette resulted in a 5.2-kb NheI DNA fragment recognized by the 3′ probe, which is diagnostic of the $Tut4^{HA\text{-}GFP}$ allele.

For the $Tut7^{HA\text{-}GFP}$ allele, we inserted the sequence encoding FLAG-HA_2-eGFP (FHG) after the endogenous ATG initiation codon within exon 2, the first coding exon of *Tut7*. A targeting construct was recombineered that contained homology arms and a loxP flanked neomycin cassette 3′ of exon 2. Southern blotting of genomic BamHI-digested DNA from individual ESC-derived clones with a 3′ probe was used to identify homologous recombinants. A 12.2-kb DNA fragment corresponds to the wild type *Tut7* locus; integration of the loxP flanked neomycin cassette 3′ of exon 2 introduced an additional BamHI site, thus decreasing the size of the NheI DNA fragment to 6.2 kb in the targeted allele. Cre-mediated recombination and excision of the loxP flanked neomycin cassette resulted in a 4.3-kb BamHI DNA fragment recognized by the 3′ probe, which is diagnostic of the $Tut7^{HA\text{-}GFP}$ allele.

For the $Tut4^{Fl}$ allele, we flanked exon 14 and 15 with loxP sites, Cre-mediated deletion of which results in out-of-frame splicing between exon 13 and 16, leading to a premature stop codon. To generate this allele, a targeting construct was generated that contains homology arms, an FRT flanked neomycin cassette and a loxP site 5′ of exon 14 and a second loxP site 3’ of exon 15. Southern blotting of genomic BamHI-digested DNA from individual ESC-derived clones with a 5′ probe was used to identify homologous recombinants. A 16.9-kb DNA fragment corresponds to the wild type *Tut4* locus; integration of the FRT flanked

neomycin cassette 3′ of exon 15 introduces an additional BamHI site, thus decreasing the size of the BamHI DNA fragment to 7.0 kb in the targeted allele. Flp-mediated recombination removed the FRT flanked neomycin cassette and generated the *Tut4*^{Fl} allele that can be identified with the 3' probe as a 13.1-kb EcoRV DNA fragment. Cre-mediated recombination and excision of exon 14 and 15 resulted in a 10.6-kb EcoRV DNA fragment recognized by the 3′ probe, which is diagnostic of the *Tut4* null (*Tut4*$^{-}$) allele.

For the *Tut7*^{Fl} allele, we flanked exon 15 with loxP sites, Cre-mediated deletion of which results in out-of-frame splicing between exon 14 and 16, leading to a premature stop codon. To generate this allele, a targeting construct was generated that contains homology arms, an FRT flanked neomycin cassette and a loxP site 5′ and 3' of exon 15. Southern blotting of genomic EcoRV-digested DNA from individual ESC-derived clones with a 5′ probe was used to identify homologous recombinants. A 15.9-kb DNA fragment corresponds to the wild type *Tut7* locus; integration of the FRT flanked neomycin cassette 5′ of exon 15 introduces an additional EcoRV site, thus decreasing the size of the EcoRV DNA fragment to 8.8 kb in the targeted allele. Flp-mediated recombination removed the FRT flanked neomycin cassette and generated the *Tut7*^{Fl} allele that can be identified with the 3' probe as a 6.9-kb EcoRV DNA fragment. Cre-mediated recombination and excision of exon 15 resulted in a 5.2-kb EcoRV DNA fragment recognized by the 3′ probe, which is diagnostic of the *Tut7* null (*Tut7*$^{-}$) allele.

In order to generate the *Tut4*^{AAD} allele we replaced wild type exon 15 with a mutant exon where the aspartic acid 1026 and 1028 codons are mutated to encode alanine. A targeting construct was recombineered that contains homology arms and a FRT flanked neomycin cassette 3′ of exon 15 that contains the D1026A and D1028A mutations. Southern blotting of genomic BamHI-digested DNA from individual ESC-derived clones with a 5′ probe was used to identify homologous recombinants. A 16.9-kb DNA fragment corresponds to the wild-type *Tut4* locus; integration of the FRT flanked neomycin cassette 3′ of exon 15 introduces an

additional BamHI as well as a PmeI site, thus decreasing the size of the BamHI DNA fragment to 12.4 kb in the targeted allele. Flp-mediated recombination and excision of the FRT flanked neomycin cassette removes the additional BamHI site and results in a 10.5-kb BamHI-PmeI DNA fragment recognized by the external 5′ probe, which is diagnostic of the *Tut4*^{AAD} allele.

The targeting for all alleles was performed in A9 ESCs. Southern blotting as described above of the individual ESC-clone-derived DNA was used to identify homologous recombinants. A9-targeted ESCs were injected into C57BL/6 eight-cell-stage embryos as described[44]. The targeted *Tut4*$^{HA\text{-}GFP\text{-}Neo/+}$ and *Tut7*$^{HA\text{-}GFP\text{-}Neo/+}$ mice were crossed to Deleter Cre mice[45] to remove the loxP flanked neomycin cassette and generate *Tut4*$^{HA\text{-}GFP/+}$ and *Tut7*$^{HA\text{-}GFP/+}$ mice, respectively. The targeted *Tut4*$^{Neo/+}$, *Tut7*$^{Neo/+}$ and *Tut4*$^{AAD\text{-}Neo/+}$ mice were crossed to FLP-expressing transgenic mice[46] to remove the FRT flanked neomycin cassette and generate *Tut4*$^{Fl/+}$, *Tut7*$^{Fl/+}$ and *Tut4*$^{AAD/+}$ mice, respectively. Mice heterozygous for the *Tut4*^{Fl} and *Tut7*^{Fl} allele were further crossed to Deleter Cre mice[45] to generate the *Tut4*$^{-}$ and *Tut7*$^{-}$ allele, respectively. The mice analyzed in this study were on a C57Bl/6 genetic background. The Zp3Cre allele[33] and the R26ERTcre allele[47] were also used in this study. To induce Cre-mediated gene deletion, mice were intraperitoneally injected with a dose of 75 mg tamoxifen/kg body weight for a total of five times. Tamoxifen was administered every other day, with a break of three weeks between the third and fourth injection. Mice were allowed to recover for at least another three weeks after the last tamoxifen injection prior to tissue collection.

All of the mice were bred and maintained at the EMBL Mouse Biology Unit, Monterotondo, in accordance with current Italian legislation (Article 9, 27. Jan 1992, number 116) under license from the Italian Health Ministry.

Southern blotting

DNA was restriction-digested, separated on a 0.8% agarose gel and transferred to a Hybond-XL membrane (Amersham) in alkaline solution (0.4 M NaOH, 1.5 M NaCl). The blot was neutralized in 2X SSC solution, UV-crosslinked at 150 mJ and then prehybridized for two hours in prehybridization solution (0.5 M Na_2HPO_4, 1 mM EDTA, 5% SDS, 3% BSA). A DNA probe was labeled with the Random Primers DNA Labeling System (Thermo Fisher Scientific) according to the manufacturer's manual. The membrane was hybridized overnight, washed in 40 mM Na_2HPO_4, 1 mM EDTA, 5% SDS and exposed on a phosphor screen (Fujifilm).

Western blotting

Cells or homogenized tissues were incubated in lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, 15% glycerol, 1 mM DTT, 0.5% sodium deoxycholate, 0.5% Triton X-100, protease inhibitors) for 10 minutes on ice. After centrifugation at 14,000 g for 10 minutes at 4°C, the supernatant was recovered, the protein extract separated on a 4-12% Bis-Tris protein gel (Invitrogen) and transferred overnight by wet transfer onto a nitrocellulose membrane (GE Healthcare). The membrane was blocked with 3% milk/0.1% Tween-20 in PBS, incubated in primary antibody in 3% milk/0.1% Tween-20 in PBS overnight, washed in 0.1% Tween-20 in PBS, and incubated with appropriate horseradish peroxidase-coupled secondary antibody (Amersham) in 3% milk/0.1% Tween-20 in PBS for one hour. Proteins were detected using the ECL Western Blotting Detection Reagent (Amersham). Antibodies used for western blotting: Anti-TUT4 (Proteintech, 18980-1-AP), anti-AGO2 (O'Carroll lab), anti-HA (Covance, MMS-101P) and anti-α-tubulin (Sigma, T9026) were used at 1:1000, anti-TUT7 (a gift from R. Pillai, University of Geneva) at 1:2000, and anti-SMC1α (Bethyl, A300-055A) at 1:10,000 for western blotting.

Immunofluorescence staining

HA-GFP-TUT4 and HA-GFP-TUT7 fusion proteins were stained using the anti-HA (Covance, MMS-101P) antibody. Ovaries were collected and fixed in 4% formaldehyde overnight and embedded in paraffin. Sections of 6 µm thickness were stained as described previously[48]. Free oocytes were fixed in 2% paraformaldehyde for 10 minutes and washed twice with 10% Normal Donkey Serum (NDS) (Sigma Aldrich), 0.1% BSA in PBS. Oocytes were then blocked with 10% NDS, 0.1 M Glycine and 2% BSA in PBS, permeabilized with 0.1% Triton X-100, 10% NDS, 0.1% BSA in PBS and washed again. The anti-β-tubulin (Sigma Aldrich, T4026) primary antibody and the Alexa Fluor Donkey (Invitrogen, A-21202) secondary antibody were used at 1:200 and 1:1000, respectively. DNA was stained with 5 µg/ml of Hoechst 33342 for 10 minutes at room temperature. Cells were mounted on slides for confocal microscopy. Images were acquired with a Leica TCS SP5 confocal microscope and Photoshop was used for cropping and other modifications that were equally performed on control and experimental samples.

Oocyte collection

For the collection of GV oocytes, three weeks old females were stimulated with 10 U of pregnant mare's serum gonadotropin (PMSG) and their ovaries were removed and placed in M2 medium. GV oocytes were released from the somatic cells via manual mechanical separation as described[4]. GV oocytes were then washed in M2 medium. For MII oocytes collection, three weeks old females were injected with 10 U of PMSG and 48 hours later with 10 U of human chorionic gonadotropin (hCG). Oocytes were collected 14 hours later from the oviduct.

RNA extraction, quantification and quality control

Total RNA was extracted using QIAzol lysis reagent (Qiagen) according to manufacturer's instructions. RNA concentration was measured with a Qubit fluorometer (Invitrogen) and quality was verified with a Bioanalyzer (Agilent Technologies).

TAIL-seq

TAIL-seq libraries were generated and analyzed as described[11] with the following modifications. Instead of performing a HiSeq asymmetric 50-250 base pair paired-end run, a 100 base pair HiSeq symmetric paired-end run was used. The starting material used for all cells and tissues (except for oocytes) was 5 µg of total RNA. Biological replicates are shown for each tissue/cell. Approximately 100-150 ng of GV oocyte total RNA was mixed with 850-900 ng of *C. elegans* total RNA in order to have sufficient RNA for TAIL-seq library preparation. To account for artifacts associated with PCR amplification, barcoded 3' adapters were used with unique molecule identifiers and only one read per ligation event was used in the final analysis. To equalize read depth, the somatic libraries were multiplexed with four libraries sequenced per lane of an Illumina HiSeq 2000, whereas each GV oocyte library was run on two lanes of the same sequencer. The original TAIL-seq pipeline[11] was used to determine the poly(A) tail length and map the reads. *C. elegans* transcripts in GV oocyte libraries were removed during the alignment of the reads to the mouse genome. The TAIL-seq learning algorithm was validated using the previously published spikes-ins (REF). The pipeline precisely determined the length of the spikes-ins (Extended Data Fig. 1a). A systematic decrease in the read recovery for the spikes-ins with longer poly(A) tails (Extended Data Fig. 1b) was observed. To account for differences in read recoveries, a linear model was fitted to the spike-ins (log2) read recovery using the spike-in's known poly(A) lengths as predictors. The model was then used to predict the relative recovery for each nucleotide in the 8 to 80 nucleotide range. For poly(A) length analyses the read recovery of

transcripts with poly(A) lengths ≤80 nucleotides was adjusted according to their predicted recovery.

Early embryo development assay

Adult females were stimulated with 10 U of PMSG. After 48 hours the females were injected with 10 U of hCG and set up with a stud C57Bl/6 male of proven fertility. Plugs were checked on the following morning and zygotes were collected in M2 media (Sigma Aldrich) from the oviduct of plugged females. Embryos were cultured under 5% CO_2 supplemented atmosphere at 37°C in KSOMaa Evolve media (Zenith Biotech) and the progression was scored every 24 hours for five days.

Oocyte *in vitro* maturation assay and NSN/SN analysis

GV oocytes were cultured in KSOMaa Evolve medium (Zenith Biotec) at 37°C and 5% CO_2 and their maturation was examined after 2, 4, 12 and 24 hours. Progression though meiosis was monitored by the germinal vesicle breakdown (GVBD) followed by the extrusion of the first polar body (PB). Cells were taken at each time point and processed as previously described for the immunofluorescence staining using anti-β-tubulin (Sigma Aldrich, T4026) antibody and Hoechst 33342 to monitor meiotic progression. For NSN/SN oocyte determination fully grown GV oocytes were collected from ovaries 44-46 hours after PMSG injection, fixed in 4% paraformaldehyde for 15 minutes and washed three times with 0.1% Tween 20, 3% BSA in PBS. Oocytes were permeabilized in 0.5% Triton X-100, 3% BSA in PBS for 35 minutes and washed three times with 0.1% Tween 20, 3% BSA in PBS. Oocytes were then stained with Hoechst 33342 for 20 minutes and mounted on Teflon coated slides (Dutscher scientific).

Gene expression analysis

For microarray profiling of oocytes, biotinylated cDNA was synthesized from total RNA using the Ovation Pico WTA System V2 kit (NuGen). The fragmentation and labeling was done using the Encore Biotin Module (NuGen). For all other cells and tissues, biotinylated cDNA was synthesized from total RNA using the Ambion WT Expression kit. The fragmentation and labeling was done using the GeneChip WT Terminal Labeling and Controls kit (Affymetrix). Next, cDNA was hybridized for 16 hours at 45°C on a GeneChip Mouse Gene 2.0 ST Array (Affymetrix). GeneChips were washed and stained in the Affymetrix Fluidics Station 450. Biological replicates were used for the analyses; GV *Tut4/7*^{CTL}, n=3; GV *Tut4/7*^{cKO}, n=4; GV *Tut4/7*^{cAAD}, n=3; Liver *Tut4/7*^{CTL}, n=3; liver *Tut4/7*^{iKO}, n=3; BM *Tut4/7*^{CTL}, n=2; BM *Tut4/7*^{iKO}, n=2; MEFs *Tut4/7*^{CTL}, n=3; MEFs *Tut4/7*^{iKO}, n=3; ESCs *Tut4/7*^{CTL}, n=3; ESCs *Tut4/7*^{iKO}, n=3).

For qRT-PCR, total RNA was reverse-transcribed using SuperScript III and random hexamers (both Invitrogen) according to manufacturer's instructions. qRT-PCR was performed using the LightCycler 480 SYBR Green I Master mix (Roche), and samples were run in technical duplicates or triplicates on a Roche LightCycler 480 instrument. C_t values were normalized against the internal control *RNA polymerase II*, *Gapdh*, or *Sod1*. Fold differences in expression levels were calculated according to the $2^{-\Delta\Delta CT}$ method[49].

Statistics

For RNA profiling, robust multi-array average (RMA) was used to normalize the raw data and differential expression was determined with the limma package[50]. Adjusted p-values for the moderated t-statistic were used. For the cluster analysis of transcripts across folliculogenesis, we used the previously described dataset[4]. Transcripts were clustered using the Markov clustering (MCL) algorithm implemented in the BioLayout express software[35].

The MCL inflation value was set at 2.2. The enrichment or depletion of upregulated genes in each cluster was determined using the hypergeometric test. Gene ontology enrichment was established with Fisher's exact test using the R topGO package. T-tests were used to compare differences between treatments assuming equal variance. All comparisons were made only using biological replicates as data points. Pearson's chi-squared test was used to determine the differences in the levels of mono and oligo-uridylation between upregulated and not upregulated genes in TUT4/7 deficient oocytes and control oocytes. Fisher's exact tests were used to evaluate the differences between oocytes in SN and NSN and follicles with developed and underdeveloped antra for *Tut4/7*^{CTL} and *Tut4/7*^{cKO} animals.

Histology

Ovaries from adult females were collected and fixed in Bouins solution overnight and embedded in paraffin. Standard Periodic Acid-Schiff (PAS) staining was performed on sections 7 μm thick.

MEF culture

Primary MEFs were derived from embryonic day 13.5 embryos according to standard protocols. Cells were cultured in MEF medium (DMEM supplemented with 12.5% fetal calf serum, 2 mM L-glutamine, 1X non-essential amino acids, 100 units/ml penicillin/streptomycin and 100 μM β-mercaptoethanol (all Gibco)) at 37°C and 7.5% CO_2. Immortalized MEFs were generated by two consecutive retroviral infections of passage 2 primary MEFs with pBabeSV40LT (a gift from G. Hannon, Cold Spring Harbor Laboratory). To induce Cre-mediated gene deletion, cells were incubated with 600 nM 4-OH-tamoxifen in MEF medium for three days. To determine the growth curve for each line, 68,000 cells per

well were plated in a 6-well plate in triplicate for each day of the experiment. Cells were trypsinized and counted with a Cellometer X2 Image Cytometer (Nexcelom).

ESC culture and differentiation

ESCs were derived from mice backcrossed one generation into the 129S2/SvPasCrl (Charles River Laboratories) background as described[51]. Cells were maintained on 0.1% gelatin in ESC medium (KnockOut DMEM supplemented with 12.5% fetal calf serum, 2 mM L-glutamine, 1X non-essential amino acids, 100 units/ml penicillin/streptomycin, 100 μM β-mercaptoethanol (all Gibco), 1 μM PD0325901, 3 μM CHIR99021 (both University of Dundee), and 20 ng/ml LIF (EMBL Heidelberg)) at 37°C and 7.5% CO_2. To induce Cre-mediated gene deletion, cells were incubated with 400 nM 4-OH-tamoxifen in ESC medium for four days. To determine the growth curve for each line, 100,000 cells per well were plated in a 6-well plate in triplicate for each day of the experiment. Cells were trypsinized and counted with a Cellometer X2 Image Cytometer (Nexcelom).

For the alkaline phosphatase staining, ESCs were seeded at 3,000 cells in a 10-cm dish and cultured for eight days. Cells were then washed with PBS, air-dried and stained with AP staining solution (100 mM Tris pH 9, 100 mM NaCl, 5 mM $MgCl_2$, 0.4 μg/ml naphtol phosphate and 1 μg/ml Fast Violet B (both Sigma)) for 15 minutes. Cells were washed again with PBS, air dried and analyzed.

Embryoid bodies (EBs) were derived by culturing ESCs according to the hanging drop method[52]. Drops of EB differentiation medium (DMEM supplemented with 20% FCS, 2 mM L-glutamine, 1X non-essential amino acids, 100 units/ml penicillin/streptomycin and 100 μM β-mercaptoethanol (all Gibco)) containing 1,000 ES cells were placed onto the lids of culture dishes and cultured for two days. EBs were then transferred onto bacterial plates and cultured

in suspension for another two days. For the cardiac differentiation assay, single EBs were plated onto gelatin-coated 24-well plates and checked daily for contractile activity.

In vitro differentiation into neural progenitor cells was achieved by culturing ESCs on 0.1% gelatin in neuronal differentiation medium (1:1 mixture of DMEM/F-12 and Neurobasal-A medium supplemented with 0.5X N-2, 0.5X B-27 (all Gibco), 25 µg/ml BSA, 1 mM L-glutamine, 10 µg/ml insulin, 150 nM thioglycerol, 1 µg/ml heparin (all Sigma) and 250 ng/ml FGF-basic (Peprotech)) for three days.

Small RNA sequencing

Small RNA libraries were generated using the NEBNext Multiplex Small RNA Library Prep Set for Illumina (Set 1) (NEB) according to manufacturer's instructions. Samples from the same tissue or cell line were multiplexed, pooled and run in one lane of HiSeq. For the analysis of the small RNA-seq data, plain miRNA counts were identified from the small RNA-seq input samples, cleaned from their 3' adapters and mapped against all mouse hairpin precursors of the miRBase, allowing up to two mismatches using Chimira[53,54]. Counts of multi-mapped reads were assigned only to the first optimal alignment call returned by BLASTn. Modification analysis was then restricted to pure modification events, i.e. mono-[nt] or poly-[nt] patterns, where [nt] can be the nucleotides U, A, C or G. Poly-[nt] modifications refer to sequences of two or more identical nucleotides. Counts from all other miRNA variants were collapsed with their respective unmodified miRNA counts. Normalization of control and experimental samples and identification of differentially expressed miRNAs was done using the DESeq2 software package[55].

44. De Fazio, S. *et al.* The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. *Nature* **480,** 259–263 (2011).

45. Schwenk, F., Baron, U. & Rajewsky, K. A cre-transgenic mouse strain for the ubiquitous deletion of loxP-flanked gene segments including deletion in germ cells. *Nucleic Acids Res.* **23,** 5080–5081 (1995).

46. Henrich, V. C. *et al.* Widespread recombinase expression using FLPeR (flipper) mice. *Genesis* **28,** 106–110 (2000).

47. Badea, T. C., Wang, Y. & Nathans, J. A noninvasive genetic/pharmacologic strategy for visualizing cell morphology and clonal relationships in the mouse. *J. Neurosci.* **23,** 2314–22 (2003).

48. Comazzetto, S. *et al.* Oligoasthenoteratozoospermia and Infertility in Mice Deficient for miR-34b/c and miR-449 Loci. *PLoS Genet.* **10,** e1004597 (2014).

49. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real- Time Quantitative PCR and the 2 ‫ ⌬⌬ C T Method. **408,** 402–408 (2001).

50. Ritchie, M. E. *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43,** e47 (2015).

51. Bryja, V. *et al.* An efficient method for the derivation of mouse embryonic stem cells. *Stem Cells* **24,** 844–849 (2006).

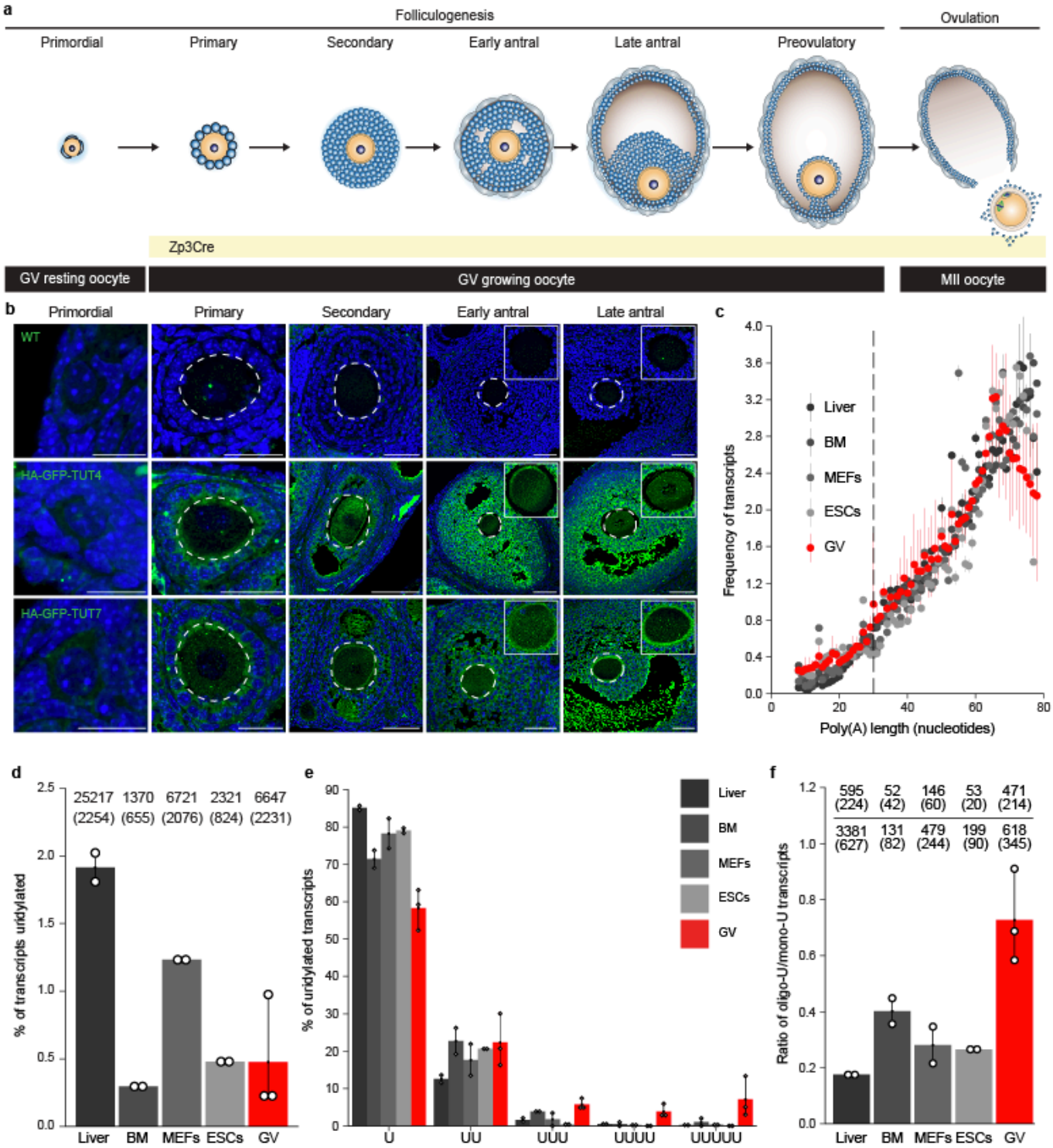
52. Keller, G. M. In vitro differentiation of embryonic stem cells. *Curr. Opin. Cell Biol.* **7,** 862–869 (1995).

53. Kozomara, A. & Griffiths-jones, S. miRBase : annotating high confidence microRNAs using deep sequencing data. **42,** 68–73 (2014).

54. Vitsios, D. M. & Enright, A. J. Sequence analysis Chimira : analysis of small RNA sequencing data and microRNA modifications. **31,** 3365–3367 (2015).

55. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. 1–21 (2014). doi:10.1186/s13059-014-0550-8

Figure 1. Morgan *et al.*



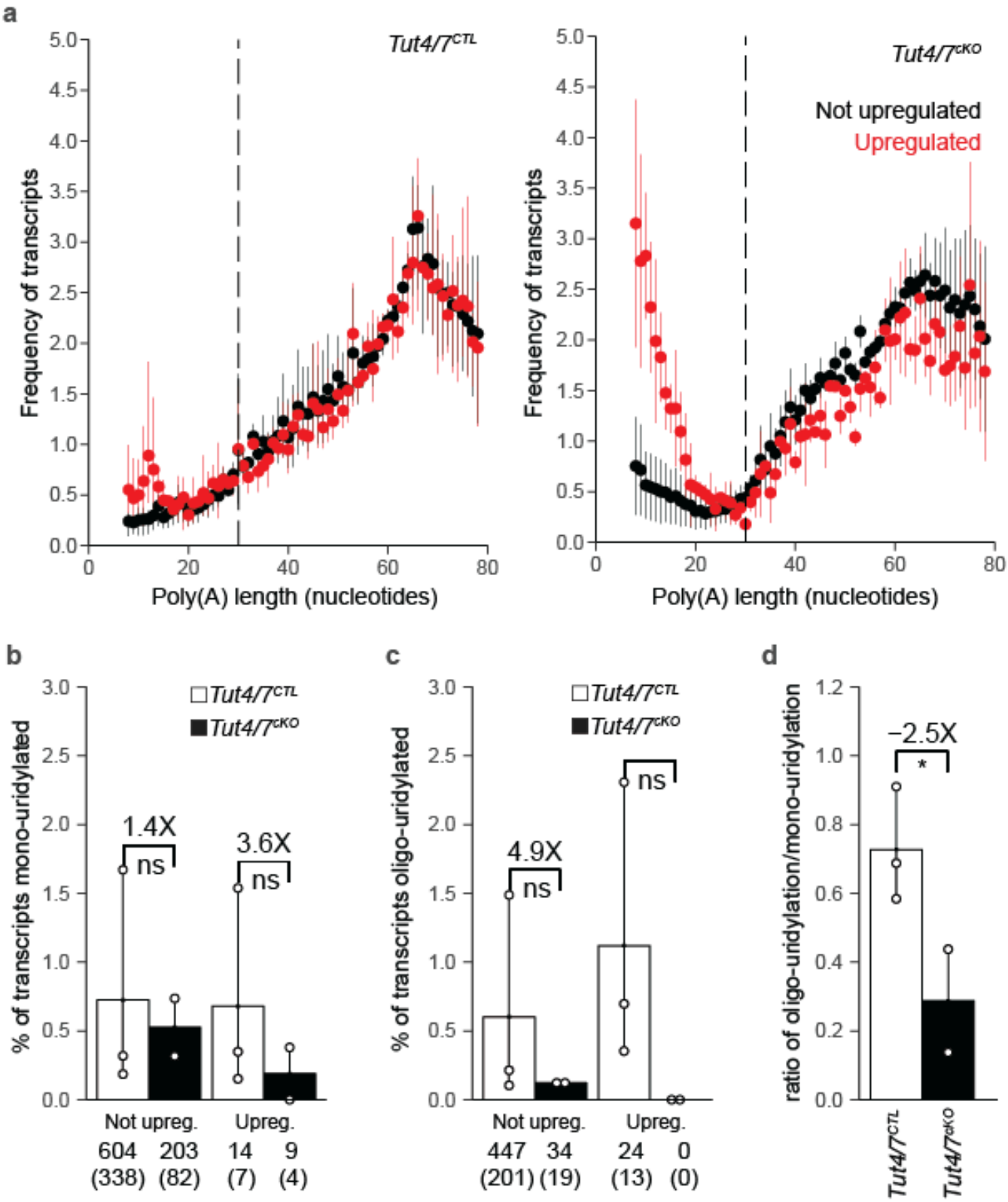
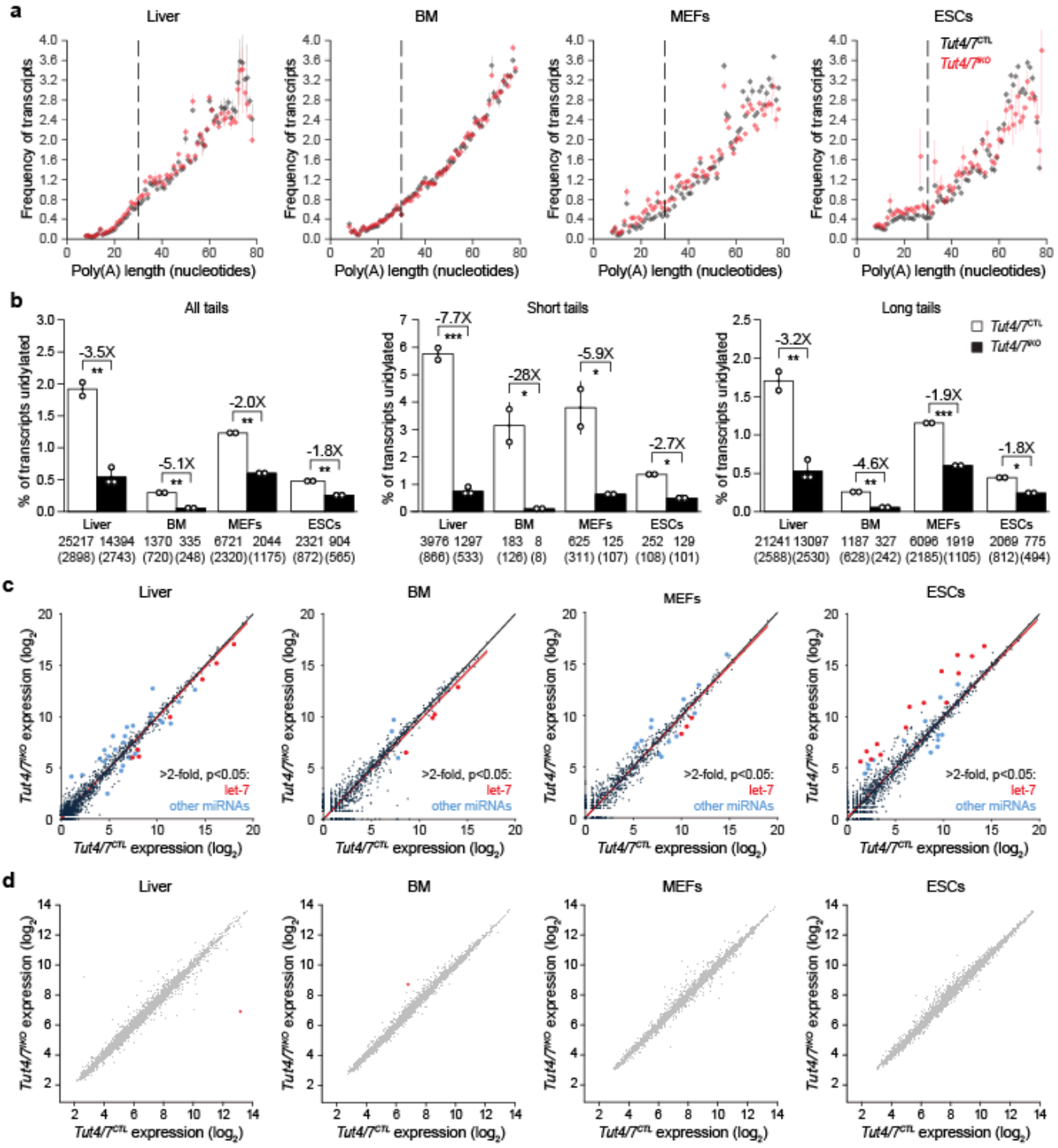

a
Tut4/7CTL
Tut4/7cKO
Not upregulated
Upregulated
Frequency of transcripts
Poly(A) length (nucleotides)
b
Tut4/7CTL
Tut4/7cKO
% of transcripts mono-uridylated
1.4X
ns
3.6X
ns
Not upreg.
Upreg.
604 203 14 9
(338) (82) (7) (4)
c
Tut4/7CTL
Tut4/7cKO
% of transcripts oligo-uridylated
4.9X
ns
ns
Not upreg.
Upreg.
447 34 24 0
(201) (19) (13) (0)
d
ratio of oligo-uridylation/mono-uridylation
−2.5X
*
Tut4/7CTL
Tut4/7cKO

Figure 5. Morgan *et al.*



Extended Data Figure 2. Morgan *et al.*