# m6A potentiates Sxl alternative pre-mRNA splicing for robust Drosophila sex determination

**Haussmann, I. , Bodi, Z. , Sanchez-Moran, E. , Mongan, N. P. , Archer, N. , Fray, R. G. and Soller, M**



**Original citation:**
Haussmann, I. , Bodi, Z. , Sanchez-Moran, E. , Mongan, N. P. , Archer, N. , Fray, R. G. and Soller, M. (2016) m6A potentiates Sxl alternative pre-mRNA splicing for robust Drosophila sex determination. Nature, volume 540 : 301-304. DOI: 10.1038/nature20577

http://dx.doi.org/10.1038/nature20577

Nature Publishing Group

# m6A potentiates *Sxl* alternative pre-mRNA splicing for robust *Drosophila* sex determination

**IRMGARD U. HAUSSMANN[1, 2], ZSUZSANNA BODI[3]*, EUGENIO SANCHEZ-MORAN[1]*, NIGEL P. MONGAN[4]*, NATHAN ARCHER[3], RUPERT G. FRAY[3] AND MATTHIAS SOLLER[1, 5]**

[1]School of Biosciences, College of Life and Environmental Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

[2]School of Life Science, Faculty of Health and Life Sciences, Coventry University, Coventry, CV1 5FB, United Kingdom

[3]School of Biosciences, Plant Science Division, University of Nottingham, Sutton Bonington, Loughborough, LE12 5RD, United Kingdom

[4]School of Veterinary Medicine and Sciences, University of Nottingham, Sutton Bonington, Loughborough, LE12 5RD, United Kingdom

Running title: m6A is required for Sxl alternative splicing

**Key Words**: mRNA methylation, *N6*-methyladenosine, alternative splicing, sex determination, dosage compensation

*equal contributing authors

[5] Corresponding author: m.soller@bham.ac.uk

**_N6_-methyladenosine (m6A) is the most common internal modification of eukaryotic messenger RNA (mRNA) and is decoded by YTH domain proteins[1-7]. The mammalian mRNA m6A methylosome is a complex of nuclear proteins that include METTL3 (Methyltransferase-like 3), METTL14, WTAP (Wilms tumour 1 associated protein) and KIAA1429. _Drosophila_ has corresponding homologues named dIME4 and dKAR4 (Inducer of meiosis-4 and Karyogamy protein-4), and Female-lethal(2)d (Fl(2)d) and Virilizer (Vir)[8-12]. In _Drosophila, fl(2)d_ and _vir_ are required for sex-dependent regulation of alternative splicing (AS) of the sex determination factor _Sex-lethal_ (_Sxl_)[13]. However, the functions of m6A in introns in the regulation of AS remain uncertain[3]. Here we show that m6A is absent in mRNA of _Drosophila_ lacking _dIME4_. In contrast to mouse and plant knock-out models[5,7,14], Drosophila _dIME4_ null mutants remain viable, though flightless and show a sex bias towards maleness. This is because m6A is required for female-specific AS of _Sxl_, which determines female physiognomy, but also translationally represses _male-specific lethal2_ (_msl-2_) to prevent dosage compensation normally occurring in males. We further show that the m6A reader protein YT521-B decodes m6A in the sex-specifically spliced intron of _Sxl_, as its absence phenocopies _dIME4_ mutants. Loss of m6A also affects AS of additional genes, predominantly in the 5'UTR, and has global impacts on the expression of metabolic genes. Requirement of m6A and its reader YT521-B for female-specific _Sxl_ AS reveal that this hitherto enigmatic mRNA modification constitutes an ancient and specific mechanism to adjust levels of gene expression.**

In mature mRNA the m6A modification is most prevalently found around the stop codon as well as in 5'UTRs and in long exons in mammals, plants and yeast[2,3,6,7,15]. Since methylosome components predominantly localize to the nucleus it has been speculated that m6A localized in

pre-mRNA introns could have a role in AS regulation in addition to such a role when present in long exons[9-12,16]. This prompted us to investigate whether m6A is required for *Sxl* AS, which determines female sex and prevents dosage compensation in females[13]. We generated a null allele of the *Drosophila METTL3* methyltransferase homologue *dIME4* by imprecise excision of a *P-element* inserted in the promoter region. The excision Δ*22-3* deletes most of the protein-coding region including the catalytic domain and is thus referred to as $dIME4^{null}$ (Fig 1a). These flies are viable and fertile, but flightless, and this phenotype can be rescued by a genomic construct restoring *dIME4* (Fig 1a and b). *dIME4* shows increased expression in the brain, and like in mammals and plants[17], localizes to the nucleus (Fig 1c,d).

Following RNAse T1 digestion and $^{32}P$ end-labeling of RNA fragments we detected m6A after G in polyA mRNA of adult flies at relatively low levels compared to other eukaryotes (m6A/A ratio: 0.06%, Fig. 1g)[2,3,5], but higher in unfertilized eggs (0.18%, Extended Data Fig. 1). After enrichment with an anti-m6A antibody m6A is readily detected in polyA mRNA, but absent from $dIME4^{null}$ (Fig. 1h-j).

As found in other systems and consistent with a potential role in translational regulation[18-21], m6A was detected in polysomal mRNA (0.1%, Fig. 1k), but not in the poly(A)-depleted ribosomal RNA (rRNA) fraction. This also confirmed that any m6A modification in rRNA is not after G in *Drosophila* (Fig. 1l).

Consistent with our hypothesis that m6A plays a role in sex determination and dosage compensation, the number of $dIME4^{null}$ females was reduced to 60% compared to the number of males ($p<0.0001$), while in the control strain female viability was 89% (Fig. 2a). The key regulator of sex determination in *Drosophila* is the RNA binding protein Sxl, which is specifically expressed in females. Sxl positively auto-regulates expression of itself and its target

*transformer* (*tra*) through AS to direct female differentiation[13]. In addition, *Sxl* suppresses translation of *msl-2* to prevent up-regulation of transcription on the X-chromosome for dosage compensation (Fig. 2b); full suppression also requires maternal factors[22]. Accordingly, female viability was reduced to 13% by removal of maternal m6A together with zygotic heterozygosity for *Sxl* and *dIME4* ($dIME4^{\Delta 22\text{-}3}$ females crossed with $Sxl^{7B0}$ males, a *Sxl* null allele, $p<0.0001$). Female viability of this genotype is completely rescued by a genomic construct (Fig. 2a) or by preventing ectopic activation of dosage compensation by removal of *msl-2* ($msl\text{-}2^{227}/Df(2L)Exel7016$, Fig. 2a). Hence, females are non-viable due to insufficient suppression of *msl-2* expression resulting in up-regulation of gene expression on the X-chromosome from reduced Sxl levels. In the absence of *msl-2*, disruption of *Sxl* AS resulted in females with sexual transformations (32%, $n$=52) displaying male-specific features such as sex combs (Fig. 2c-e), which were mosaic to various degrees indicating that Sxl threshold levels are affected early during establishment of sexual identities of cells and/or their lineages[13]. In the presence of maternal dIME4, *Sxl* and *dIME4* do not genetically interact ($Sxl^{7B0}/FM7$ females crossed with $dIME4^{null}$ males, 103% female viability, $n$=118). In addition, Sxl is required for germline differentiation in females and its absence results in tumorous ovaries[23]. Consistent with this we detected tumorous ovaries in $Sxl^{7B0}$/+; $dIME4^{null}$/+ daughters from $dIME4^{null}$ females (22%, $n$=18, Extended Data Fig. 2), but not in homozygous $dIME4^{null}$ or heterozygous $Sxl^{7B0}$ females ($n$=20 each).

Furthermore, levels of the *Sxl* female-specific splice form were reduced to ~50% consistent with a role for m6A in *Sxl* AS (Fig. 2f and Extended Data Fig. 3a). As a result, female-specific splice forms of *tra* and *msl-2* were also significantly reduced in adult females (Fig. 2f and Extended Data Fig. 3b,c).

To obtain more comprehensive insights into *Sxl* AS defects in *dIME4*$^{null}$ females, we examined splice junction reads from RNA-seq. Besides the significant increase in inclusion of the male-specific *Sxl* exon in *dIME4*$^{null}$ females (Fig. 2f- h, and Extended Data Fig. 3a), cryptic splice sites and increased numbers of intronic reads were detected in the regulated intron. Consistent with our RT-PCR analysis of *tra*, the reduction of female splicing in the RNA sequencing is modest, and as a consequence, AS differences of Tra targets *dsx* and *fru* were not detected in whole flies, suggesting cell-type specific fine-tuning required to generate splicing robustness rather than being an obligatory regulator (Extended Data Fig. 4a-c). In agreement with dosage compensation defects as main consequence of *Sxl* miss-regulation in *dIME4*$^{null}$ mutants, X-linked, but not autosomal, genes are significantly up-regulated in *dIME4*$^{null}$ females compared to the control ($p<0.0001$, Extended Data Fig. 4d,e).

Further, we also find enrichment of *Sxl* mRNA in pull-downs with an m6A antibody compared to m6A-deficient yeast mRNA added for quantification (Fig. 2i). This enrichment is comparable to what was observed for m6A-methylated mRNA in yeast[24].

To further map m6A sites in the intron of *Sxl* we employed an in vitro m6A methylation assay using *Drosophila* nuclear extracts and labeled substrate RNA. m6A methylation activity was detected in the vicinity of alternatively spliced exons (Fig. 2j, RNAs B, C, and E). Further fine-mapping localized m6A in RNAs C and E to the proximity of Sxl binding sites (Extended Data Fig. 5). Likewise, the female-lethal single amino acid substitution alleles *fl(2)d*$^{1}$ and *vir*$^{2F}$ interfere with Sxl recruitment, resulting in impaired *Sxl* auto-regulation and inclusion of the male-specific exon[25]. Female lethality of these alleles can be rescued by *dIME4*$^{null}$ heterozygosity ($p<0.0001$, Fig. 2k), further demonstrating involvement of the m6A methylosome in *Sxl* AS.

Next, we globally analyzed AS changes in *dIME4*$^{null}$ females compared to the wild-type control strain. As described earlier (Fig. 2h), a statistically significant reduction in female-specific AS of *Sxl* ($\Delta$psi=0.34, q=9x10$^{-8}$) was observed. In addition, 243 AS events in 163 genes were significantly different in *dIME4*$^{null}$ females (q<0.05, $\Delta$psi>0.2), equivalent to ~2% of alternatively spliced genes in *Drosophila* (Suppl. Table 1). Six genes for which the AS products could be distinguished on agarose gels were confirmed by RT-PCR (Extended Data Fig. 6). Interestingly, lack of *dIME4* did not affect global AS and no specific type of AS event was preferentially affected. However, alternative first exons (18% vs 33%) and mutually exclusive exon (2% vs 15%) events were reduced mostly to the extent of retained introns (16% vs 6%), alternative donor (16% vs 9%) and unclassified events (14% vs 6%) compared to a global breakdown of AS in *Drosophila* (Extended Data Fig. 7a). Interestingly, the majority of affected AS events in *dIME4*$^{null}$ were located to the 5'UTR, and these genes had a significantly higher number of AUGs in their 5'UTR compared to the 5'UTRs of all genes (Extended Data Fig. 7b,c). Such feature had been shown relevant to translational control under stress conditions[26].

The majority of the 163 differentially alternatively spliced genes in dIME4 females are broadly expressed (59%), while most of the remainder are expressed in the nervous system (33%), consistent with higher expression of *dIME4* in this tissue (Extended Data Fig. 7d). Accordingly, gene ontology (GO) analysis revealed a highly significant enrichment for genes in synaptic transmission (p<7x10$^{7}$, Suppl. Table 1).

Since the absence of m6A affects AS, m6A marks are probably deposited co-transcriptionally before splicing. Co-staining of polytene chromosomes with antibodies against HA-tagged dIME4 and RNA Pol II revealed broad co-localization of dIME4 with sites of transcription (Fig. 3a-e), but not with condensed chromatin visualized with antibodies against histone H4 (Fig. 3f-i).

Furthermore, localization of dIME4 to sites of transcription is RNA-dependent, as staining for dIME4, but not for RNA Pol II, was reduced in an RNase-dependent manner (Fig. 3j,k).

Although m6A levels after G are low in *Drosophila* compared to other eukaryotes, broad co-localization of dIME4 to sites of transcription suggests profound effects on the gene expression landscape. Indeed, differential gene expression analysis revealed 408 differentially expressed genes (≥2-fold change, q≤0.01) where 234 genes were significantly up- and 174 significantly down-regulated in neuron-enriched head/thorax of adult $dIME4^{null}$ females (q<0.01, at least two-fold, Suppl. Table 2). Cataloguing these genes according to function reveals prominent effects on gene networks involved in metabolism including reduced expression of 17 genes involved in oxidative phosphorylation (p<0.0001, Suppl. Table 2). Notably, overexpression of the m6A mRNA demethylase FTO in mice leads to an imbalance in energy metabolism resulting in obesity[27].

Next, we tested whether either of the two substantially divergent YTH proteins, YT521-B and CG6422 (Fig. 4a) decodes m6A marks in *Sxl* mRNA. When transiently transfected into male S2 cells, YT521-B localizes to the nucleus, whereas CG6422 is cytoplasmic (Fig. 4b-d, Ext. Data Fig. 8). Nuclear YT521-B can switch *Sxl* AS to the female mode and also binds to the *Sxl* intron in S2 cells (Fig. 4e,f). In vitro binding assays with the YTH domain of YT521-B indeed demonstrate increased binding of m6A-containing RNA (Ext. Data Fig. 9). In vivo, YT521-B also localizes to sites of transcription (Ext. Data Fig. 10).

To further examine the role of YT521-B in decoding m6A we analyzed *Drosophila* strain $YT521\text{-}B^{MI02006}$ where a transposon in the first intron disrupts *YT521-B*. This allele is also viable ($YT521\text{-}B^{MI02006}/Df(3L)Exel6094$; Fig. 4g,h,j), and phenocopies the flightless phenotype and the female *Sxl* splicing defect of $dIME4^{null}$ (Fig. 4h,i). Likewise, removal of maternal *YT521-B*

together with zygotic heterozygosity for *Sxl* and *YT521-B* reduced female viability ($p<0.0001$, Fig. 4j) and resulted in sexual transformations (57%, $n=32$) such as male abdominal pigmentation (Fig. 4k-m). In addition, overexpression of YT521-B results in male lethality, which can be rescued by removal of *dIME4* further reiterating the role of m6A in *Sxl* AS ($p<0.0001$, Fig. 4n). Since *YT521-B* phenocopies *dIME4* for *Sxl* splicing regulation it is the main nuclear factor for decoding m6A present in the proximity of the Sxl binding sites. YT521-B bound to m6A assists Sxl in repressing inclusion of the male-specific exon, thus providing robustness to this vital gene regulatory switch (Fig. 4o).

Nuclear localization of m6A methylosome components suggested a role for this "fifth" nucleotide in AS regulation. Our discovery of the requirement of m6A and its reader YT521-B for female-specific *Sxl* AS has important implications for understanding the fundamental biological function of this enigmatic mRNA modification. Its key role in providing robustness to *Sxl* AS to prevent ectopic dosage compensation and female lethality, together with localization of the core methylosome component dIME4 to sites of transcription, indicates that the m6A modification is part of an ancient, yet unexplored mechanism to adjust gene expression. Hence, the recently reported role of m6A methylosome components in human dosage compensation[28,29] further support such role and suggests that m6A-mediated adjustment of gene expression might be a key step to allow for development of the diverse sex determination mechanisms found in nature.

## Acknowledgments

We thank J. Horabin, N. Perrimon and Bloomington, Harvard and Kyoto stock centers for fly lines, BacPAc for DNA clones, E. Zaharieva and M.L. Li for help with imaging, W. Arlt and R.

Michell for comments on the manuscript, and J-Y. Roignant for communication of results prior to publication. We acknowledge funding from the BBSRC (BB/M008606/1) to RF.

**Author contributions**

I.U.H. and M.S. performed biochemistry, cell biology and genetic experiments, E.S.M. stained chromosomes, and Z.B., N.A. and R.F. performed biochemistry experiments. N.M. analyzed sequencing data. I.U.H., R.F. and M.S. conceived the project and wrote the manuscript with help from N.M. and Z.B.

**Author information**

 Correspondence and requests for materials should be addressed to M.S. (m.soller@bham.ac.uk).

## References:

1 Luo, S. & Tong, L. Molecular basis for the recognition of methylated adenines in RNA by the eukaryotic YTH domain. *Proc Natl Acad Sci U S A* **111**, 13834-13839 (2014).

2 Meyer, K. D. *et al.* Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635-1646 (2012).

3 Dominissini, D. *et al.* Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**, 201-206 (2012).

4 Perry, R. P. & Kelley, D. E. Existence of methylated messenger RNA in mouse L cells. *Cell* **1**, 37-42 (1974).

[5] Zhong, S. *et al.* MTA is an Arabidopsis messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. *Plant Cell* **20**, 1278-1288 (2008).

[6] Schwartz, S. *et al.* High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. *Cell* **155**, 1409-1421 (2013).

[7] Ke, S. *et al.* A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev* **29**, 2037-2053 (2015).

[8] Liu, J. *et al.* A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol* **10**, 93-95 (2013).

[9] Horiuchi, K. *et al.* Identification of Wilms' tumor 1-associating protein complex and its role in alternative splicing and the cell cycle. *J Biol Chem* **288**, 33292-33302 (2013).

[10] Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G. & Rottman, F. M. Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA* **3**, 1233-1247 (1997).

[11] Penalva, L. O. *et al.* The Drosophila fl(2)d gene, required for female-specific splicing of Sxl and tra pre-mRNAs, encodes a novel nuclear protein with a HQ-rich domain. *Genetics* **155**, 129-139 (2000).

[12] Niessen, M., Schneiter, R. & Nothiger, R. Molecular identification of virilizer, a gene required for the expression of the sex-determining gene Sex-lethal in Drosophila melanogaster. *Genetics* **157**, 679-688 (2001).

[13] Schutt, C. & Nothiger, R. Structure, function and evolution of sex-determining systems in Dipteran insects. *Development* **127**, 667-677 (2000).

14 Geula, S. *et al.* Stem cells. m6A mRNA methylation facilitates resolution of naive pluripotency toward differentiation. *Science* **347**, 1002-1006 (2015).

15 Luo, G. Z. *et al.* Unique features of the m6A methylome in Arabidopsis thaliana. *Nat Commun* **5**, 5630 (2014).

16 Xiao, W. *et al.* Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. *Mol Cell* **61**, 507-519 (2016).

17 Hongay, C. F. & Orr-Weaver, T. L. Drosophila Inducer of MEiosis 4 (IME4) is required for Notch signaling during oogenesis. *Proc Natl Acad Sci U S A* **108**, 14855-14860 (2011).

18 Bodi, Z., Bottley, A., Archer, N., May, S. T. & Fray, R. G. Yeast m6A Methylated mRNAs Are Enriched on Translating Ribosomes during Meiosis, and under Rapamycin Treatment. *PLoS One* **10**, e0132090 (2015).

19 Wang, X. *et al.* N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* **161**, 1388-1399 (2015).

20 Meyer, K. D. *et al.* 5' UTR m(6)A Promotes Cap-Independent Translation. *Cell* **163**, 999-1010 (2015).

21 Zhou, J. *et al.* Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature* **526**, 591-594 (2015).

22 Zaharieva, E., Haussmann, I. U., Brauer, U. & Soller, M. Concentration and localization of co-expressed ELAV/Hu proteins control specificity of mRNA processing. *Mol Cell Biol* **35**, 3104-3115 (2015).

23 Salz, H. K. Sex, stem cells and tumors in the Drosophila ovary. *Fly (Austin)* **7**, 3-7 (2012).

[24] Bodi, Z., Button, J. D., Grierson, D. & Fray, R. G. Yeast targets for mRNA methylation. *Nucleic Acids Res* **38**, 5327-5335 (2010).

[25] Hilfiker, A., Amrein, H., Dubendorfer, A., Schneiter, R. & Nothiger, R. The gene virilizer is required for female-specific splicing controlled by Sxl, the master gene for sexual development in Drosophila. *Development* **121**, 4017-4026 (1995).

[26] Starck, S. R. *et al.* Translation from the 5' untranslated region shapes the integrated stress response. *Science* **351**, aad3867 (2016).

[27] Church, C. *et al.* Overexpression of Fto leads to increased food intake and results in obesity. *Nat Genet* **42**, 1086-1092 (2010).

[28] Moindrot, B. *et al.* A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. *Cell Rep* **12**, 562-572 (2015).

[29] Patil, D. P. *et al.* m6A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* **537**, 369-373 (2016).

## Figure legends

**Figure 1: Analysis of *dIME4* null mutants and m6A methylation in *Drosophila*.** **a**, Genomic organization of the *dIME4* locus depicting the transposon (black triangle) used to generate the deletion Δ*22-3*, which is a *dIME4* null allele and the hemaglutinin (HA)-tagged genomic rescue fragment. **b**, Flight ability of $dIME4^{null}/Df(3R)Exel6197$ shown as mean±SE ($n$=3). *gdIME4*: genomic rescue construct. **c** and **d**, Nuclear localization of dIME4::HA in eye discs and brain neurons expressed from *UAS*. Scale bars: 50 and 1 µm. **e**, Schematic diagram of a 2D thin layer chromatography (TLC). **f**, TLC from an *in vitro* transcript containing m6A. **g**, TLC from mRNA of adult flies. **h** and **i**, TLC of fragmented mRNA after enrichment with an anti-m6A antibody

from wild type (**h**) and *dIME4*$^{null}$ (**i**, overexposed). **j**, Quantification of immunoprecipitated $^{32}$P label shown as normalized mean (*n*=2). **k** and **l**, TLC from mRNA (**k**) or rRNA (**l**) from polysomes from wild-type flies.

**Figure 2: m6A methylation is required for *Sex-lethal* AS in sex determination and dosage compensation.** **a**, Female viability of indicated genotypes devoid of maternal m6A (*n*: total number of flies). **b**, Schematic depicting Sxl control of female differentiation. **c-e**, Front legs of indicated genotypes. Scale bar: 100 μm. The arrowhead points towards the position of the sex comb normally present only in males. **f**, Ratio of sex-specific splice isoforms from adult females from RT-PCR shown as mean±SE (*n*=3, p<0.01). **g**, RT-PCR for male-specific *Sxl* splicing in control and *dIME4*$^{null}$ females. **h**, Sashimi plot depicting Tophat-mapped RNA sequencing reads and exon junction reads from control and *dIME4*$^{null}$ females below the annotated gene model. Male-specific splice junction reads are circled and cryptic splice sites are boxed. RNA fragments used for m6A *in vitro* methylation assays are indicated at the bottom. **i**, Presence of m6A in *Sxl* transcripts detected by m6A immunoprecipitation followed by qPCR from nuclear mRNA of early embryos (shown as mean, *n*=2). **j**, 1D-TLC of *in vitro* methylated, [$^{32}$P]-ATP-labeled substrate RNAs shown in **g**. Nucleotide markers from *in vitro* transcripts in the absence (M1) or presence (M2) of m6A. The right part shows an overexposure of the same TLC. **k**, Rescue of female lethality of female-lethal *fl(2)d*$^{1}$ and *vir*$^{2F}$ alleles by removal of one copy of *dIME4*.

**Figure 3: dIME4 co-localizes to sites of transcription.** **a-e**, Polytene chromosomes from salivary glands expressing dIME::HA stained with anti-Pol II (red, **c**), anti-HA (green, **d**) and DNA (DAPI, blue, **e**), or merged (yellow, **a** and **b**). **f-i**, Polytene chromosomes stained with anti-

Pol II (red, **h**), anti-histone H4 (green, **g**) and DNA (blue, **i**), or merged (yellow, **f**). Polytene chromosomes treated with low (**j**, 2 μg/ml) and high (**k**, 10 μg/ml) RNase A concentration prior to staining with anti-Pol II, anti-histone H4 and DNA. Scale bars in **a, j** and **k** are 20 μm and in **e** and **i** are 5 μm.

**Fig 4: YTH protein YT521-B decodes m6A methylation in *Sxl*.** **a**, Domain organization of *Drosophila* YTH proteins (YTH domain in green). n: nuclear, c: cytoplasmic **b-d**, Cellular localization and size of HA-tagged YT521-B and CG6422 in S2 cells. Scale bar: 1 μm. **e**, Suppression of male-specific *Sxl* AS upon expression of Sxl and YT521-B, but not CG6422 in male S2 cells. **f**, Binding of YT521-B to pre-mRNA of the regulated *Sxl* intron. **g**, Genomic organization of the *YT521-B* locus depicting the transposon (black triangle) disrupting the ORF. **h**, Flight ability of $YT521\text{-}B^{MI02006}/Df(3L)Exel6094$ shown as mean±SE ($n$=3). **i**, *Sxl* AS in female wild-type and $YT521\text{-}B^{MI02006}/Df(3L)Exel6094$ flies. **j**, Female viability of indicated genotypes ($n$: total number of flies) reared at 29º C. **k-m**, Abdominal pigmentation of indicated genotypes reared at 29 ºC. The arrowhead points towards the position of the dark pigmentation normally present only in males. Scale bar: 100 μm. **n**, YT521-B was overexpressed from a *UAS* transgene with *tubulinGAL4* ($2^{nd}$) in wild type or $dIME4^{null}$ at 27 ºC. **o,** Model for female-specific *Sxl* AS by Sxl, m6A and its reader YT521-B in co-operatively suppressing inclusion of the male-specific exon.

## Online Methods

### Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

**_Drosophila_ genetics, generation of constructs and transgenic lines**

The deletion allele *dIME4*$^{\Delta 22\text{-}3}$ was obtained from imprecise excision of the transposon *P{SUPor-P}KrT95D* and mapped by primers 5933 F1 (CTCGCTCTATTTCTCTTCAGCACTCG) and 5933 R9 (CCTCCGCAACGATCACATCGCAATCGAG). To obtain a viable line of *dIME4*$^{null}$, the genetic background was cleaned by out-crossing to *Df(3R)Exel6197*. Flight ability was scored as number of flies capable of flying out of a petri-dish within 30 sec for groups of 15-20 flies for indicated genotypes. Viability was calculated from the numbers of females compared to males of the correct genotype and statistical significance was determined by a $\chi^2$ test (GraphPad Prism). Unfertilized eggs were generated by expressing sex-peptide in virgin females as described[30].

The genomic rescue construct was retrieved by recombineering (Genebridges) from BAC clone *CH321-79E18* by first cloning homology arms with *Spe*I and *Acc*65I into *pUC3GLA* separated by an *EcoR*V site for linearization (CTCCGCCGCCGGAACCGCCGCCTCCTCCGCCACTTTGCAGGTTGAGCGGACCGCCT CCAGGGCCGCTGCCGCCGGTGCCGCTGATATCCCAGCATGGTAGCTGCGGCCACTCC TAGTCCCGCCTTTAACCACAGCTTGGGGTCCTCCGTCATCAGGCCGAATTGCCTCGA G). An HA-tag was then fused to the end of the ORF using two PCR amplicons and *Sac*I and *Xho*I restriction sites. This construct was the inserted into *PBac{y+-attB-3B}VK00002* at 76A as described[31].

346 The dIME4 UAS construct was generated by cloning the ORF from fly cDNA into a modified
347 *pUAST* with primers Adh dMT-A70 F1 EI
348 (GCAGAATTCGAGATCtAAAGAGCCTGCTAAAGCAAAAAAGAAGTCACCATGGCAGA
349 TGCGTGGGACATAAAATCAC) and dMT-A70 HA R1 Spe
350 (GGTAACTAGTCTTTTGTATTCCATTGATCGACGCCGCATTGG) by adding a translation
351 initiation site from the *Adh* gene and two copies of an HA tag to the end of the ORF. This
352 construct was then also inserted into *PBac{y+-attB-3B}VK00002* at 76A.
353 For transient transfection in S2 cells, *YT52B-1* and *CG6422* ORFs were amplified from fly
354 cDNA by a combination of nested and fusion PCR incorporating a translation initiation site from
355 the *Adh* gene using primers CG6422 adh F1
356 (GCCTGCTAAAGCAAAAAAGAAGTCACCACATGTCAGGCGTGGATCAGATGAAAAT
357 ACCAG), pact adh CG6422 F1
358 (CCAGAGACCCCGGATCCAGATATCAAAGAGCCTGCTAAAGCAAAAAAGAAGTCAC
359 CAC), CG6422 adh R1, (GATTCCTGCGAACAGGTCCCGTGGGCGAAAC) and CG6422 3'
360 F1 (CCCACGGGACCTGTTCGCAGGAATCTAG), CG6422 3' R1
361 (CATTGCTTCGCATTTTATCCTTGTCCGTGTCCTTAAAGCGCACGCCGATTTTAATTTG
362 ), pact CG6422 3xHA R1
363 (GTGGAGATCCATGGTGGCGGAGCTCGAGGAATATTCATTGCTTCGCATTTTATCCTT
364 GTC) for *CG6422* and primers YT521 adh F1,
365 (AAGCAAAAAAGAAGTCACATGCCAAGAGCAGCCCGTAAACAAACGCTGCCGATGC
366 GCGAG), pact adh YT521 F1
367 (CCAGAGACCCCGGATCCAGATATCAAAGAGCCTGCTAAAGCAAAAAAGAAGTCAC
368 ATGCC), YT521 adh R1

(TGCCATCCGGGCGAATCCTGCAAATTTACCACTCTCGTTGACCGAGAAAATGAGCA GGAC) and YT521 3' F1(GCAGGATTCGCCCGGATGGCAGCCCCCTCAC), Pact YT521 R1 (GGTGGAGATCCATGGTGGCGGAGCTCGAGCGCCTGTTGTCCCGATAGCTTCGCTG) for *YT521-B*, and cloned into a modified *pACT* using Gibson Assembly (NEB) also incorporating HA epitope tags at the C terminus. Constructs were verified by Sanger sequencing. The Sxl-HA expression vector was a gift from N. Perrimon[32].

The YT521-B UAS construct was generated by sub-cloning the ORF from the pACT vector into a modified *pUAST* with primers YT521 adh F1 (AAGCAAAAAAGAAGTCACATGCCAAGAGCAGCCCGTAAACAAACGCTGCCGATGC GCGAG), YT521 adh F2 (TAGGGAATTGGGAATTCGAGATCTAAAGAGCCTGCTAAAGCAAAAAAGAAGTCAC ATGCC) and YT521 3' R1 (GGGCACGTCGTAGGGGTACAGACTAGTCTCGAGGCGCCTGTTGTCCCGATAGCTTC GCTG) by adding a translation initiation site from the *Adh* gene and two copies of an HA tag to the end of the ORF. This construct was then also inserted into *PBac{y+-attB-3B}VK00002* at 76A.

Essential parts of all DNA constructs were sequence verified.

**Cell culture, transfections and immune-staining of S2 cells**

S2 cells (ATCC) were cultured in Insect Express medium (Lonza) with 10% heat-inactivated FCS and 1% penicillin/streptomycin. The *Drosophila* S2 cell line was verified to be male by analysing *Sxl* alternative splicing using species-specific primers Sxl F2 (ATGTACGGCAACAATAATCCGGGTAG) and Sxl R2

(CATTGTAACCACGACGCGACGATG) to confirm species and gender (Ext. Data Fig 8). Transient transfections were done with Mirus Reagent (Bioline) according to the manufacturer's instruction and cells were assayed 48 h after transfection for protein expression or RNA binding of expressed proteins. To adhere S2 cells to a solid support, Concanavalin A (Sigma) coated glass slides (in 0.5 mg/ml) were added 1 d prior to transfection, and cells were stained 48 h after transfection with antibodies as described. Transfections and follow up experiments were repeated at least once.

**RNA extraction, RT-PCR, qPCR, immune-precipitations and Western blots**

Total RNA was extracted using Tri-reagent (SIGMA) and reverse transcription was done with Superscript II (Invitrogen) according to the manufacturer's instructions using an oligodT17V primer. PCR for *Sxl, tra, msl2* and *ewg* was done for 30 cycles with 1 µl of cDNA with primers Sxl F2, Sxl R2 or Sxl NP R3 (GAGAATGGGACATCCCAAATCCACG), Sxl M F1 (GCCCAGAAAGAAGCAGCCACCATTATCAC), Sxl M R1 (GCGTTTCGTTGGCGAGGAGACCATGGG), tra FOR (GGATGCCGACAGCAGTGGAAC), tra REV (GATCTGGAGCGAGTGCGTCTG), msl-2 F1 (CACTGCGGTCACACTGGCTTCGCTCAG), msl-2 R1 (CTCCTGGGCTAGTTACCTGCAATTCCTC), ewg 4F and ewg 5R and quantified with ImageQuant (BioRad)[22]. Experiments included at least three biological replicates.

For qPCR reverse transcription was carried out on input and pull-down samples spiked with yeast RNA using ProtoScript II reverse transcriptase and random nanomers (NEB). Quantitative PCR was carried out using 2x SensiMix Plus SYBR Low ROX master mix (Quantace) using normalizer primers ACT1 F1 (TTACGTCGCCTTGGACTTCG) and ACT1 R1

(TACCGGCAGATTCCAAACCC) and for Sxl, Sxl ZB F1 (CACCACAATGGCAGCAGTAG) and Sxl ZB R1 (GGGGTTGCTGTTTGTTGAGT). Samples were run in triplicate for technical repeats and duplicate for biological repeats. Relative enrichment levels were determined by comparison with yeast *ACT1*, using the $2^{-\Delta\Delta C'_T}$ method[33].

For immunoprecipitations of *Sxl* RNA bound to Sxl or YTH proteins, S2 cells were fixed in PBS containing 1% formaldehyde for 15 min, quenched in 100 mM glycine and disrupted in IP-Buffer (150 mM NaCl, 50 mM Tris–HCL, pH 7.5, 1% NP-40, 5% glycerol). After IP with anti-HA beads (Sigma) for 2 h in the presence of Complete protein inhibitors (Roche) and 40 U RNase inhibitors (Roche), IP precipitates were processed for *Sxl* RT-PCR using gene-specifc RT primer SP NP2 (CATTCCGGATGGCAGAGAATGGGAC) and PCR primers Sxl NP intF (GAGGGTCAGTCTAAGTTATATTCG) and Sxl NP R3 as described[31]. Western blots were done as described using rat anti-HA (1:50, clone 3F10, Roche) and HRP coupled secondary goat anti-rat antibodies (Molecular Probes)[34]. All experiments were repeated at least once from biological samples.

**Analysis of m6A levels**

PolyA mRNA from at least two rounds of oligo dT selection was prepared according to the manufacturer (Promega). For each sample, 10-50 ng of mRNA was digested with 1 µl of Ribonuclease T1 (1000 U/µl; Fermentas) in a final volume of 10 µl in polynucleotide kinase buffer (PNK, NEB) for 1 h at 37 ºC. The 5' end of the T1-digested mRNA fragments were then labeled using 10 U T4 PNK (NEB) and 1 µl [$\gamma$-$^{32}$P]-ATP (6000 Ci/mmol; Perkin-Elmer). The labeled RNA was precipitated, resuspended in 10 µl of 50 mM sodium acetate buffer (pH 5.5), and digested with P1 nuclease (Sigma-Aldrich) for 1 h at 37 ºC. Two microliters of each sample

was loaded on cellulose TLC plates (20x20 cm; Fluka) and run in a solvent system of isobutyric acid: 0.5 M $NH_4OH$ (5:3, v/v), as first dimension, and isopropanol:HCl:water (70:15:15, v/v/v), as the second dimension. TLCs were repeated from biological replicates. The identification of the nucleotide spots was carried out using m6A-containing synthetic RNA. Quantification of $^{32}P$ was done by scintillation counting (Packard Tri-Carb 2300TR). For the quantification of spot intensities on TLCs or gels, a storage phosphor screen (K-Screen; Kodak) and Molecular Imager FX in combination with QuantityOne software (BioRad) were used.

For immunoprecipitation of m6A mRNA, polyA mRNA was digested with RNase T1 and 5' labeled. The volume was then increased to 500 µl with IP buffer (150 mM NaCl, 50 mM Tris–HCL, pH 7.5, 0.05% NP-40). IPs were then done with 2 µl of affinity-purified polyclonal rabbit m6A antibody (Synaptic Systems) and protein A/G beads (SantaCruz).

**Polysome profiles**

Whole fly extracts were prepared from 20-30 adult *Drosophila* previously frozen in liquid $N_2$ and ground into fine powder in liquid $N_2$. Cells were then lysed in 0.5 ml lysis buffer (0.3 M NaCl, 15 mM $MgCl_2$, 15 mM Tris-HCl pH 7.5, cycloheximide 100 µg/ml, heparin (sodium salt) 1 mg/ml, 1% Triton X-100). Lysates were loaded on 12 ml sucrose gradients and spun for two h at 38 000 rpm at 4 $^{o}$C. After the gradient centrifugation 1 ml fractions were collected and precipitated in equal volume of isopropanol. After several washes with 80% ethanol the samples were resuspended in water and processed. Experiments were done in duplicate.

**Nuclear extract preparation and *in vitro* m6A methylation essays**

460 *Drosophila* nuclear extracts were prepared from Kc cells as described[35]. Templates for *in vitro*
461 transcripts were amplified from genomic DNA using the primers listed below and *in vitro*
462 transcribed with T7 polymerase in the presence of [α-$^{32}$P]-ATP. DNA templates and free
463 nucleotides were removed by DNase I digestion and Probequant G-50 spin columns (GE
464 Healthcare), respectively. Markers were generated by using *in vitro* transcripts with or without
465 m6ATP (Jena Bioscience), which were then digested with RNase T1, kinased with PNK in the
466 presence of [γ-$^{32}$P]-ATP. After phenol extraction and ethanol precipitation, transcripts were
467 digested to single nucleotides with P1 nuclease as above. For *in vitro* methylation, transcripts
468 (0.5-1x$10^6$ cpm) were incubated for 45 min at 27 ºC in 10 µl containing 20 mM potassium
469 glutamate, 2 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, 0.5 mM S-adenosylmethionine disulfate
470 tosylate (Abcam), 7.5% PEG 8000, 20 U RNase protector (Roche) and 40% nuclear extract.
471 After phenol extraction and ethanol precipitation, transcripts were digested to single nucleotides
472 with P1 nuclease as above, and then separated on cellulose F TLC plates (Merck) in 70%
473 ethanol, previously soaked in 0.4 M $MgSO_4$ and dried[36]. *In vitro* methylation assays were done
474 from biological replicates at least in duplicates.
475 Primers to amplify parts of the *Sxl* alternatively spliced intron from genomic DNA for *in vitro*
476 transcription with T7 polymerase were Sxl A T7 F
477 (GGAGCTAATACGACTCACTATAGGGAGAGGATATGTACGGCAACAATAATCCGGGT
478 AG) and Sxl A R (CGCAGACGACGATCAGCTGATTCAAAGTGAAAG), Sxl B T7 F
479 (GGAGCTAATACGACTCACTATAGGGAGAGCGCTCGCATTTATCCCACAGTCGCAC)
480 and Sxl B R (GGGTGCCCTCTGTGGCTGCTCTGTTTAC), Sxl C T7 F
481 (GGAGCTAATACGACTCACTATAGGGGTCGTATAATTTATGGCACATTATTCAG) and
482 Sxl C R (GGGAGTTTTGGTTCTTGTTTATGAGTTGGGTG), Sxl D T7 F

(GGAGCTAATACGACTCACTATAGGGAGAAAACTTCCAGCCCACACAACACACAC ) and Sxl D R (GCATATCATATTCGGTTCATACATTTAGGTCTAAG), Sxl E T7 F (GGAGCTAATACGACTCACTATAGGGAGAGGGGAAGCAGCTCGTTGTAAAATAC) and Sxl E R (GATGTGACGATTTTGCAGTTTCTCGACG), Sxl F T7 F (GGAGCTAATACGACTCACTATAGGGAGAGGGGGATCGTTTTGAGGGTCAGTCTAAG ) and Sxl NP2, Sxl C T7 F and Sxl C1 R (GTAGTTTTGCTCGGCATTTTATGACCTTGAGC), Sxl C2 F (GGAGCTAATACGACTCACTATAGGGAGACTCTCATTCTCTATATCCCTGTGCTGACC ) and Sxl C2 R (CTAATTTCGTGAGCTTGATTTCATTTTGCACAG), Sxl C3 F (GGAGCTAATACGACTCACTATAGGGAGACTGTGCAAAATGAAATCAAGCTCACGA AATTAG) and Sxl C R, Sxl E T7 F and Sxl E1 R (AAAAAAATCAAAAAAATAATCACTTTTGGCACTTTTTCATCAC), Sxl E2 F (GGAGCTAATACGACTCACTATAGGGAGATGAAAAAGTGCCAAAAGTGATTATTTTT TTG), Sxl E2 R (AAAAGCATGATGTATTTTTTTTTTTTTGTACTTTCGAATCACCG), Sxl E3 F (GGAGCTAATACGACTCACTATAGGGAGACGGTGATTCGAAAGTACAAAAAAAAAA AAATAC) and Sxl E R, Sxl C4 F (GAGCTAATACGACTCACTATAGGGAGAAATACTAAAACATCAAACCGCAAGCAGA GCAGC) and Sxl C4 R (GAGTGCCACTTCAAAATCTCAGATATGC), Sxl C5 F (CTAATACGACTCACTATAGGGAGACTCTTTTTTTTTTCTTTTTTTTACTGTGCAAAA TG) and Sxl C5 R (AAAAAAATATGCAAAAAAAAAAAGGTAGGGCACAAAGTTCTCAATTAC), Sxl C6 F (GAGCTAATACGACTCACTATAGGGAGACTGTGCAAAATGAAATCAAGCTCACGAA

ATTAG) and Sxl C6 R (CAATTTCACTATATGTACGAAAACAAAAGTGAG), Sxl E4 F (GGAGCTAATACGACTCACTATAGGGAGAACCAAAATTCGACGTGGGAAGAAAC) and Sxl E4 R (TAATCACTTTTGGCACTTTTTCATCACATTAAC), Sxl E5 F (GGCTAATACGACTCACTATAGGGAGATTTTTTTGATTTTTTTAAAGTGAAAATGTGC TCC) and Sxl E5 R (CACCGAAAAAAAATAAAAAAAAAATAATCATGGGACTATACTAG), Sxl E6 F (GGCTAATACGACTCACTATAGGGAGACTTAAGTGCCAATATTTAAAGTGAAACCAA TTG) and Sxl E6 R (CCCCCAGTTATATTCAACCGTGAAATTCTGC).

**Illumina sequencing and analysis of differential gene expression and AS**

Total RNA was extracted from 15 pulverized head/thoraces previously flash frozen in liquid nitrogen, using Trizol reagent from *white* (*w*) control and *w*; *dIME4*$^{\Delta 22\text{-}3}$ females that have been outcrossed for several generations to *w*; *Df(3R)Exel6197* to equilibrate genetic background. Total RNA was treated with DNase I (Ambion) and stranded libraries for Illumina sequencing were prepared after polyA selection from total RNA (1 μg) with the TruSeq stranded mRNA kit (Illumina) using random primers for reverse transcription according to the manufacturer's instructions. Pooled indexed libraries were sequenced on an Illumina HiSeq2500 to yield 40-46 million paired-end 100 bp reads, and in a second experiment 14-19 million single-end 125 bp reads for three controls and mutants each. After demultiplexing, sequence reads were aligned to the *Drosophila* genome (dmel-r6.02) using Tophat2.0.6[37]. Differential gene expression was determined by Cufflinks-Cuffdiff and the FDR-correction for multiple testings to raw P values with $q<0.05$ considered significant[38]. AS was analysed by SPANKI[39] and validated for selected genes based on length differences detectable on agarose gels. Illumina sequencing, differential

gene expression and AS analysis was done by Fasteris (Switzerland). For dosage compensation analysis, differential expression analysis of X-linked genes versus autosomal genes in *dIME4*$^{null}$ mutant was done by filtering Cuffdiff data by a p value expression difference significance of $p<0.05$, which corresponds to a false discovery rate of 0.167 to detect subtle differences in expression consistent with dosage compensation. Visualization of sequence reads on gene models and splice junctions reads in Sashimi plots was done using Integrated Genome Viewer[40]. For validation of AS by RT-PCR as described above, the following primers were used: Gprk2 F1 (CCAACCAGCCGAAACTCACAGTGAAGC) and Gprk2 R1 (CAGGGTCTCGGTTTCAGACACAGGCGTC), fl(2)d F1 (GCAGCAAACGAGAAATCAGCTCGCAGCGCAG) and fl(2)d R1 (CACATAGTCCTGGAATTCTTGCTCCTTG), A2bp1 F3 (CTGTGGGGCTCAGGGGCATTTTTCCTTCCTC) and A2bp1 R1 (CTCCTCTCCCGTGTGTCTTGCCACTCAAC), cv-c F1 (GGGTTTCCACCTCGACCGGGAAAAGTCG) and cv-c R1 (GCGTTTGCGGTTGCTGCTCGCGAAGAGAG), CG8312 F1 (GCGCGTGGCCTCCTTCTTATCGGCAGTC) and CG8312 R1 (GCGTGGCCACTATAAAGTCCACCTCATC), Chas F2 (CCGATTCGATTCGATTCGATCCTCTCTTC) and Chas R1 (GTCGGTGTCCTCGGTGGTGTTGGTGGAG). GO enrichment analysis was done with FlyMine. For the analysis of uATGs, a custom R script was used to count the uATGs in 5'UTRs in all ENSEMBL isoforms of those genes which are differentially spliced in dIME4 mutants, that were then compared to the mean number of ATGs in all *Drosophila* ENSEMBL 5'UTRs using a t-test. Gene expression data were obtained from flybase.

**Custom R Script**

```
> fasta_file <-read.fasta("Soller_UTRs.fa", as.string=T)# read fasta file
> pattern <-"atg" # the pattern to look for
> dict <-PDict(pattern, max.mismatch = 0)#make a dictionary of the pattern to look for
> seq <- DNAStringSet( unlist(fasta_file)[1:638])#make the DNAstrinset from the DNAsequences ie all 638 UTRs related to the 156 genes identified in spanki
> result <-vcountPDict(dict,seq)#count the pattern in each of the sequences
> write.csv2(result, "result.csv")

> fasta_file <-read.fasta("dmel-all-five_prime_UTR-r6.07.fa", as.string=T)# read fasta file
> pattern <-"atg" # the pattern to look for
> dict <-PDict(pattern, max.mismatch = 0)#make a dictionary of the pattern to look for
> seq <- DNAStringSet( unlist(fasta_file)[1:29822])#make the DNAstrinset from the DNAsequences ie all UTRs
> result <-vcountPDict(dict,seq)#count the pattern in each of the sequences
> write.csv2(result, "result_allutrs.csv")
```

**Polytene chromosome preparations and stainings**

dIME4 or YT521-B were expressed in salivary glands with *C155-GAL4* from a *UAS* transgene. Larvae were grown at 18 ºC under non-crowded conditions. Salivary glands were dissected in PBS containing 4% formaldehyde and 1% Triton X-100, and fixed for 5 min, and then for another 2 min in 50% acetic acid containing 4% formaldehyde, before placing them in lactoacetic acid (lactic acid:water:acetic acid, 1:2:3). Chromosomes were then spread under a

siliconized cover slip and the cover slip removed after freezing. Chromosome were blocked in PBT containing 0.2% BSA and 5% goat serum and sequentially incubated with primary antibodies (mouse anti-PolII H5, 1:1000, Abcam, or rabbit anti-histone H4, 1:200, Santa-Cruz, and rat anti-HA MAb 3F10, 1:50, Roche) followed by incubation with Alexa488- and/or Alexa647-coupled secondary antibodies (Molecular Probes) including DAPI (1 µg/ml, Sigma). RNase A treatment (4 and 200 µg/ml) was done before fixation for 5 min. Ovaries were analyzed as previously described[41].

**RNA binding assays**

The YTH domain (aa 207-423) was PCR amplified with oligos YTHdom F1 (CAGGGGCCCCTGTCGACTAGTCCCGGGAATGGTGGCGGCAACGGCCG) and R1 (CACGATGAATTGCGGCCGCTCTAGATTACTTGTAGATCACGTGTATACCTTTTTCTC GC) and cloned with Gibson assembly (NEB) into a modified pGEX expression vector to express a GST-tagged fusion protein. The YTH domain was cleaved while GST was bound to beads using Precession protease. Electrophoretic mobility shift assays and UV cross-linking assays were performed as described[35,42]. Quantification was done using ImageQuant (BioRad) by measuring free RNA substrate to calculate bound RNA from input. All binding assays were done at least in triplicates.

**Data availability statement:** RNA-seq data that support the findings of this study have been deposited at GEO under the accession number GSE79000 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE79000), combining the single-end (GSE78999) and paired-end (GSE78992) experiments

598 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE78999 and
599 http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE78992, respectively). All other data
600 generated or analysed during this study are included in this published article (and its
601 Supplementary Information files).
602
603 **Additional references**

604 [30] Haussmann, I. U., Hemani, Y., Wijesekera, T., Dauwalder, B. & Soller, M. Multiple
605 pathways mediate the sex-peptide-regulated switch in female Drosophila reproductive
606 behaviours. *Proc Biol Sci* **280**, 20131938 (2013).
607 [31] Haussmann, I. U., Li, M. & Soller, M. ELAV-mediated 3'-end processing of ewg
608 transcripts is evolutionarily conserved despite sequence degeneration of the ELAV-
609 binding site. *Genetics* **189**, 97-107 (2011).
610 [32] Yan, D. & Perrimon, N. spenito is required for sex determination in Drosophila
611 melanogaster. *Proc Natl Acad Sci U S A* **112**, 11606-11611 (2015).
612 [33] Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-
613 time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408
614 (2001).
615 [34] Haussmann, I. U., White, K. & Soller, M. Erect wing regulates synaptic growth in
616 Drosophila by integration of multiple signaling pathways. *Genome Biol* **9**, R73 (2008).
617 [35] Soller, M. & White, K. ELAV inhibits 3'-end processing to promote neural splicing of
618 ewg pre-mRNA. *Genes Dev* **17**, 2526-2538 (2003).

[36] Harper, J. E., Miceli, S. M., Roberts, R. J. & Manley, J. L. Sequence specificity of the human mRNA N6-adenosine methylase in vitro. *Nucleic Acids Res* **18**, 5735-5741 (1990).

[37] Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**, R36 (2013).

[38] Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7**, 562-578 (2012).

[39] Sturgill, D. *et al.* Design of RNA splicing analysis null models for post hoc filtering of Drosophila head RNA-Seq data with the splicing analysis kit (Spanki). *BMC Bioinformatics* **14**, 320 (2013).

[40] Robinson, J. T. *et al.* Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26 (2011).

[41] Soller, M., Bownes, M. & Kubli, E. Control of oocyte maturation in sexually mature Drosophila females. *Dev Biol* **208**, 337-351 (1999).

[42] Soller, M. & White, K. ELAV multimerizes on conserved AU4-6 motifs important for ewg splicing regulation. *Mol Cell Biol* **25**, 7580-7591 (2005).

## Extended Data figure legends

**Extended Data Figure 1: m6A levels in unfertilized eggs. a** and **b,** TLC from maternal total RNA (**a**) and mRNA (**b**) present in unfertilized eggs. The arrow indicates m6A.

**Extended Data Figure 2: dIME4 supports Sxl in directing germline differentiation. a-c,** Representative ovarioles of wild type (**a**), $dIME4^{null}/dIME4^{null}$ (**b**) and *Sxl/+;* $dIME4^{null}/+$ (**c**), and a tumerous ovary of a *Sxl/+;* $dIME4^{null}/+$ female (**d**). The tumorous ovary consisting mostly of

undifferentiated germ cells in (**d**) is indicated with a bracket and the oviduct with an asterisk. The scale bar in (**d**) is 100 μm.

**Extended Data Figure 3: *dIME4* is required for female-specific splicing of *Sxl*, *tra* and *msl-2*. a-c,** RT-PCR of *Sxl* (**a**), *tra* (**b**) and *msl-2* (**c**) sex-specific splicing in wild-type males and females, and $dIME4^{null}$ males and females. 100 bp markers are shown on the left.

**Extended Data Figure 4: AS of sex determination genes and differential expression of X-linked genes in $dIME4^{null}$ females. a-c,** Sashimi plot depicting Tophat-mapped RNA sequencing reads and exon junction reads below the annotated gene model for sex-specific AS of *tra*, *fru* and *dsx*. The thickness of lines connecting splice junctions corresponds to the number of junction reads also shown. ss: splice site. **d,** Significantly ($p<0.05$, $q<0.166853$) differentially expressed gene expression values expressed as reads per kb of transcript per million mapped reads (RPKM) were +1 log transformed and Spearman r correlation values determined for X-linked and autosomal genes in wild-type and $dIME4^{null}$ *Drosophila.* **e,** The proportion of autosomal and X-linked genes that were significantly either up- or down-regulated in $dIME4^{null}$ as compared to wild-type *Drosophila* were statistically compared using $\chi^2$ with Yates' continuity correction. GraphPad Prism was used for statistical comparisons. Similar results as for the single-read RNA-seq experiment were obtained for the pair-end RNA sequencing experiment.

**Extended Data Figure 5: m6A methylation sites map to the vicinity of *Sxl* binding sites. a,** Schematic of the *Sxl* alternatively-spliced intron around the male specific exon depicting substrate RNAs used for *in vitro* m6A methylation. Solid lines depict fragments containing m6A

methylation and dashed lines fragments where m6A was absent. **b** and **c,** 1D-TLC of *in vitro* methylated [$^{32}$P]-ATP-labeled substrate RNAs shown in (**a**). Markers are *in vitro* transcripts in the absence (M1) or presence (M2) of m6A $^{32}$P-labeled after RNase T1 digestion. The right part in (**b**) and (**c**) shows an overexposure of the same TLC.

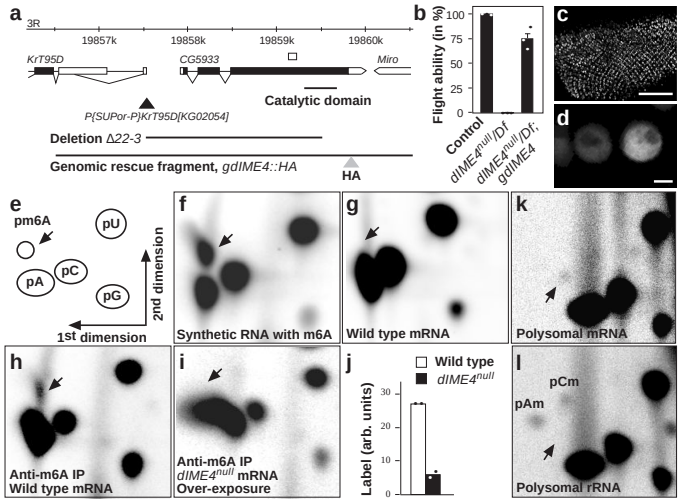
**Extended Data Figure 6: RT-PCR validation of differential AS in *dIME4$^{null}$*. a-f,** Sashimi plot depicting Tophat-mapped RNA sequencing reads and exon junction reads below the annotated gene model of indicated genes on the left, and RT-PCR of AS shown on the right using primers depicted on top. The thickness of lines connecting splice junctions corresponds to the number of junction reads also shown.

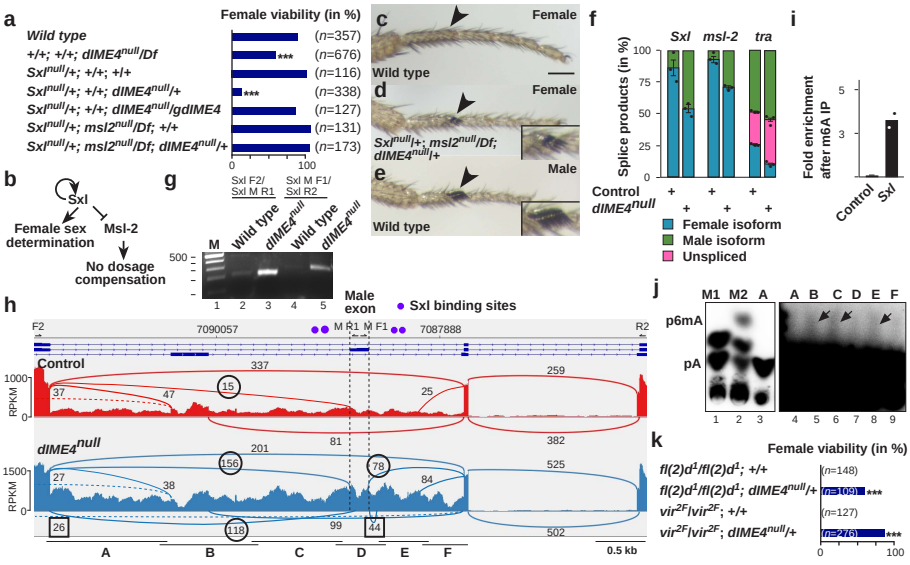
**Extended Data Figure 7: *dIME4* affects AS predominantly in 5'UTRs in genes with a higher than avarage number of upstream AUGs. a** and **b,** Classification of differential AS in *dIME4$^{null}$* according to splicing event (**a**) and location of the event in the mRNA (**b**). **c**, Quantification of upstream AUGs in all annotated 5'UTRs (white) or in alternative isoforms differentially spliced between wild type and *dIME4$^{null}$*. All *Drosophila* UTRs were accessed in fasta format from Flybase (version r6.07), (ftp://ftp.flybase.net/genomes/Drosophila_melanogaster/current/fasta/). A custom R script was used to count the number of ATG sequences in all *Drosophila* 5′UTRs and from the genes identified by the Spanki analysis comprising 638 5′UTRs. A *T* test then used to statistically compare the number of ATGs present in the 638 5′UTRs of the differentially-spliced genes as compared to all 29822 *Drosophila* 5′UTRs. **d** and **e,** Classification of differentially alternative spliced genes in *dIME4$^{null}$* according to expression pattern (**d**) or function (**e**).

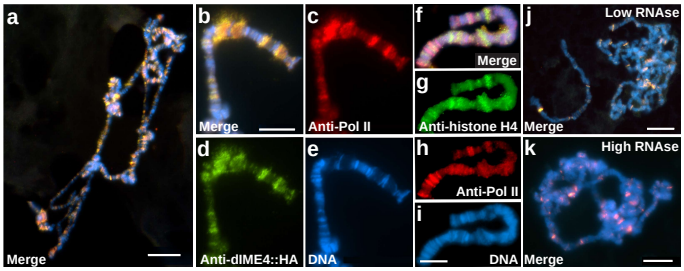
**Extended Data Figure 8: *Drosophila* S2 cells are male.** RT-PCR of *Sxl* AS in females, males and S2 cells. 100 bp markers are shown on the left.

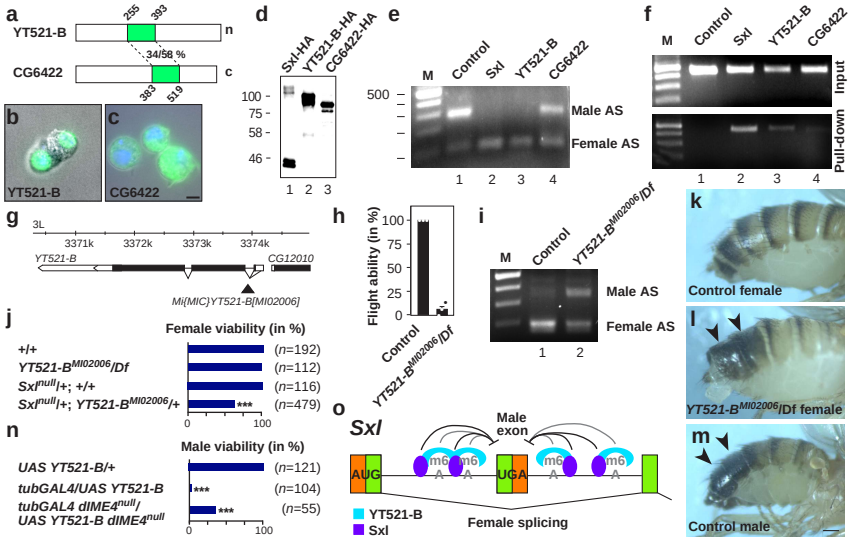
**Extended Data Figure 9: Preferential binding of the YTH domain of YT521-B to m6A-containing RNA. a,** Coomassie-stained gel depicting the recombinant YTH domain (aa 207-423) of YT521-B. **b** and **c,** Electrophoretic mobility shift assay of YTH domain binding to *Sxl* RNA fragment C with or without m6A (50%) and quantification of RNA bound to the YTH domain shown as mean±SE ($n$=3). Note that the YTH domain does not form a stable complex with RNA (asterisk) and that this complex falls apart during the run or forms aggregates in the well. **d,** In solution UV crosslinking of the YTH domain to *Sxl* RNA fragment C at 0.25 μM, 1 μM, 4 μM and 16 μM (lanes 1-4).

**Extended Data Figure 10: YT521-B co-localizes to sites of transcription. a-d**, Polytene chromosomes from salivary glands expressing YT521-B::HA stained with anti-Pol II (red, **b**), anti-HA (green, **c**) and DNA (DAPI, blue, **d**), or merged (yellow, **a**). Scale bars are 5 μm.


a
3R
19857k
19858k
19859k
19860k
KrT95D
CG5933
Miro
Catalytic domain
P[SUPor-P]KrT95D[KG02054]
Deletion Δ22-3
Genomic rescue fragment, gdIME4::HA
HA
b
Flight ability (in %)
100
75
50
25
0
Control
dIME4null/Df
dIME4null/Df; gdIME4
c
d
e
pm6A
pU
pC
pA
pG
2nd dimension
1st dimension
f
Synthetic RNA with m6A
g
Wild type mRNA
k
Polysomal mRNA
h
Anti-m6A IP
Wild type mRNA
i
Anti-m6A IP
dIME4null mRNA
Over-exposure
j
Wild type
dIME4null
Label (arb. units)
30
20
10
0
l
pCm
pAm
Polysomal rRNA


a
Female viability (in %)
Wild type (n=357)
+/+; +/+; dIME4null/Df ***
Sxlnull/+; +/+; +/+ (n=116)
Sxlnull/+; +/+; dIME4null/+ ***
Sxlnull/+; +/+; dIME4null/gdIME4
Sxlnull/+; msl2null/Df; +/+
Sxlnull/+; msl2null/Df; dIME4null/+
0 100
b
Sxl
Female sex determination
Msl-2
No dosage compensation
g
Sxl F2/ Sxl M R1
Sxl M F1/ Sxl R2
M
Wild type
dIME4null
Wild type
dIME4null
500
1 2 3 4 5
c
Female
Wild type
d
Female
Sxlnull/+; msl2null/Df; dIME4null/+
e
Male
Wild type
f
Sxl msl-2 tra
Splice products (in %)
100 75 50 25 0
Control
dIME4null
Female isoform
Male isoform
Unspliced
i
Fold enrichment after m6A IP
5 3
Control Sxl
h
Male exon
Sxl binding sites
F2 7090057 M R1 M F1 7087888 R2
Control
1000 RPKM 0
37 47 15 337 25 259 81 382
dIME4null
1500 RPKM 0
27 38 156 201 78 84 525
26 118 99 44 502
A B C D E F
0.5 kb
j
M1 M2 A
A B C D E F
p6mA
pA
1 2 3 4 5 6 7 8 9
k
Female viability (in %)
fl(2)d1/fl(2)d1; +/+ (n=148)
fl(2)d1/fl(2)d1; dIME4null/+ ***
vir2F/vir2F; +/+ (n=127)
vir2F/vir2F; dIME4null/+ ***
0 100


a
Merge
b
Merge
c
Anti-Pol II
f
Merge
g
Anti-histone H4
j
Low RNAse
Merge
d
Anti-dIME4::HA
e
DNA
h
Anti-Pol II
i
DNA
k
High RNAse
Merge


a
YT521-B
CG6422
255
393
34/58 %
383
519
b
YT521-B
c
CG6422
d
Sxl-HA
YT521-B-HA
CG6422-HA
100
75
58
46
1 2 3
e
M
Control
Sxl
YT521-B
CG6422
500
Male AS
Female AS
1 2 3 4
f
M
Control
Sxl
YT521-B
CG6422
Input
Pull-down
1 2 3 4
g
3L
3371k
3372k
3373k
3374k
YT521-B
CG12010
Mi{MIC}YT521-B[MI02006]
h
Flight ability (in %)
100
75
50
25
0
Control
YT521-BMI02006/Df
i
M
Control
YT521-BMI02006/Df
Male AS
Female AS
1 2
j
Female viability (in %)
+/+ (n=192)
YT521-BMI02006/Df (n=112)
Sxlnull/+; +/+ (n=116)
Sxlnull/+; YT521-BMI02006/+ *** (n=479)
0
100
n
Male viability (in %)
UAS YT521-B/+ (n=121)
tubGAL4/UAS YT521-B *** (n=104)
tubGAL4 dIME4null/ UAS YT521-B dIME4null *** (n=55)
0
100
o
Sxl
Male exon
AUG
UGA
m6A
YT521-B
Sxl
Female splicing
k
Control female
l
YT521-BMI02006/Df female
m
Control male