The tiny, conserved zinc-finger protein GTSF1 helps PIWI proteins achieve their full catalytic potential

Amena Arif[1,2], Shannon Bailey[2], Natsuko Izumi[3], Todd A. Anzelon[4], Deniz M. Ozata[2,6], Cecilia Andersson[2], Ildar Gainetdinov[2], Ian J. MacRae[4], Yukihide Tomari[3,5] & Phillip D. Zamore[2,*]

[1]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Chan Medical School, Worcester, MA 01605, USA

[2]Howard Hughes Medical Institute and RNA Therapeutics Institute, University of Massachusetts Chan Medical School, 368 Plantation Street, Worcester, MA 01605, USA

[3]Laboratory of RNA Function, Institute for Quantitative Biosciences, The University of Tokyo, Tokyo, Japan

[4]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

[5]Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Japan

[6]Current address: Department of Molecular Biosciences, Stockholm University, SE-106 91 Stockholm, Sweden

*Correspondence: phillip.zamore@umassmed.edu (P.D.Z.)

Abstract

Argonaute proteins use nucleic acid guides to find and bind specific DNA or RNA target sequences. Argonaute proteins can be found in all kingdoms of life, and play diverse biological functions including genome defense, gene regulation, and chromosome partitioning. Many Argonautes retain their ancestral endoribonuclease activity, cleaving the phosphodiester bond between target nucleotides t10 and t11. In animals, a specialized class of Argonautes, the PIWI proteins, use 21–35 nt PIWI-interacting RNAs (piRNAs) to direct transposon silencing, protect the germline genome, and regulate gene expression during gametogenesis[1]. The piRNA pathway is required for fertility in one or both sexes of nearly all animals. Both piRNA production and function require RNA cleavage catalyzed by PIWI proteins. Spermatogenesis in mice and other placental mammals requires three distinct, developmentally regulated PIWI proteins: MIWI (PIWIL1), MILI (PIWIL2), and MIWI2 (PIWIL4)[2–4]. The piRNA-guided endoribonuclease activities of MIWI and MILI are essential to produce functional sperm[5,6]. piRNA-directed silencing in mice and insects also requires Gametocyte-Specific Factor 1 (GTSF1), a PIWI-associated protein of unknown function[7–12]. Here, we report that GTSF1 potentiates the weak, intrinsic, piRNA-directed RNA cleavage activities of PIWI proteins, transforming them into efficient endoribonucleases. GTSF1 represents the first example of an auxiliary protein that potentiates the catalytic activity of an Argonaute protein.

Main

In animals, 21–35 nt PIWI-interacting RNAs (piRNAs) direct PIWI proteins to silence transposons and regulate gene expression[1,13–27]. Invertebrates produce piRNAs and PIWI proteins in both the soma and the germline[28–35] but mammalian piRNAs act only in the germline[36–43]. Mice lacking any of their three PIWI proteins—MIWI2 (PIWIL4), MILI (PIWIL2), and MIWI (PIWIL1)—or other proteins required for piRNA production are invariably male sterile[1–4]. As in other animals, mouse piRNA production requires catalytically active PIWI proteins, MILI and MIWI[5,6]. Transposon silencing is the ancestral function of piRNAs; uniquely, placental mammals also produce pachytene piRNAs, which first appear shortly after the onset of meiosis I (refs. 40–42,44–46) and reach a peak abundance in spermatocytes rivalling that of ribosomes[16]. Pachytene piRNAs tune the abundance of mRNAs required for spermiogenesis, the process by which round spermatids become sperm[47–50]. Pachytene piRNAs have been proposed to direct MIWI and MILI to cleave specific mRNAs, ensuring appropriate levels of their protein products[47,50–52]. Testing this idea has been thwarted by the absence of a cell-free system in which MIWI or MILI can be loaded with synthetic piRNAs of defined sequence.

Recombinant MIWI loaded with a piRNA of defined sequence

We used lentiviral transduction to engineer a stable HEK293T cell line over-expressing epitope-tagged MIWI. Tagged MIWI was captured from cell lysate using anti-FLAG antibody coupled to paramagnetic beads, incubated with a synthetic piRNA bearing a monophosphorylated 5′ terminus and 2′-*O*-methylated 3′ end, and eluted from the magnetic beads using 3XFLAG peptide. MIWI loaded with a synthetic piRNA (MIWI piRISC; Fig. 1a), but not unloaded apo-MIWI (Extended Data Fig. 1a), cleaved a 5′ ^{32}P-radiolabled target RNA fully complementary to the synthetic guide (Extended Data Fig. 1b, c and Supplementary Data Fig. 1a).

Recombinant MIWI bound stably to an RNA guide bearing a 5′ monophosphate but not to a guide with a 5′ hydroxyl group (Extended Data Fig. 1d). The MID domain of Argonaute proteins contains a 5′ monophosphate-binding pocket that anchors the RNA guide to the protein[53–59]. Mutations predicted to disrupt 5′ monophosphate-binding perturb PIWI function[55,60–64]. We immobilized MIWI on paramagnetic beads via its 3XFLAG tag, incubated it with guide RNA, washed the beads, eluted the MIWI piRISC with 3XFLAG tag peptide, and tested its ability to cleave a fully complementary target RNA. Incubation with a 5′ monophosphorylated but not an otherwise identical 5′ hydroxy guide yielded MIWI piRISC that cleaved target RNA (Extended Data Fig. 1d). In vivo, the methyltransferase HENMT1 adds a 2′-*O*-methyl group to the 3′ ends of piRNAs. *Henmt1*$^{-/-}$ male mice and *henn1*$^{-/-}$ worms are sterile[65–71]. Terminal 2′-*O*-methyl modification likely stabilizes small RNAs against degradation by cellular ribonucleases rather than secures the guide to the protein[72,73]. Consistent with this function for 2′-*O*-methylation, piRNAs bearing a 3′ terminal 2′ hydroxyl or 2′-*O*-methyl were equally functional in directing target cleavage, provided that the piRNA was 5′ monophosphorylated (Extended Data Fig. 1d).

Purified MIWI RISC is a sluggish enzyme

Although affinity-purified MIWI loaded with a piRNA bearing 5′ monophosphorylated and 3′, 2′-*O*-methylated termini specifically cleaved a fully complementary target RNA at the phosphodiester bond that links target nucleotides t10 to t11, the rate of cleavage (0.01 min^{-1}) was >300 times slower than that catalyzed by mouse AGO2 RISC (≥ 3 min^{-1})[74,75]. At physiological temperature (37°C), mouse AGO2 RISC catalyzes multiple rounds of target cleavage[76]. By contrast, 5 nM MIWI piRISC cleaved only ~15% of the target RNA (5 nM) after 1 h (Fig. 1b, Extended Data Fig. 1e, and Supplementary Data Fig. 1b). Inefficient target cleavage by MIWI piRISC was not caused by the presence of the amino-terminal 3XFLAG-SNAP tandem tag: removing the tandem tags

using tobacco etch virus protease generated an untagged protein (Extended Data Fig. 1f and Supplementary Data Fig. 1c) whose target cleavage kinetics were identical to that of piRISC assembled with the tagged MIWI (Extended Data Fig. 1g).

The ubiquitously expressed arginine methyltransferase PRMT5 modifies the amino terminus of PIWI proteins, allowing it to bind Tudor domain-containing proteins, many of which are required for piRNA biogenesis, gametogenesis, and fertility[77–79]. A potential explanation for the sluggish activity of recombinant MIWI piRISC is that it lacks arginine methylation. We used mass spectrometry to map the positions of methyl arginine in our affinity-purified MIWI. All of the arginine residues previously shown to be methylated in endogenous MIWI immunoprecipitated from mouse testis[77] were methylated in recombinant MIWI produced in HEK293T cells (Extended Data Fig. 1h).

Another possible explanation for the inefficiency of target cleavage by MIWI piRISC is that the recombinant protein, although properly arginine methylated, lacks other post-translational modifications. To test this idea, we immobilized apo-MIWI, incubated it with wild-type (C57BL/6) mouse testis lysate in the presence of an ATP regenerating system at 25°C for 15 min, removed the testis lysate by washing, and then loaded MIWI with a synthetic piRNA and eluted the resulting piRISC. Pre-incubation of MIWI with testis lysate either before or after loading with a piRNA failed to increase its target cleaving activity (Extended Data Fig. 2a, b). We conclude that neither a missing post-translational modification nor a tightly associated protein partner is likely to explain the low activity of MIWI piRISC compared to AGO2 RISC.

MIWI piRISC requires an auxiliary factor to efficiently cleave its target

Adding testis lysate increased the rate of target cleavage by affinity-purified MIWI piRISC ~20-fold (Fig. 1b; Extended Data Fig. 2b). This effect cannot be attributed to the lysate contributing additional piRISC, because testis lysate alone failed to generate detectable cleavage product (Fig. 1b).

118 MIWI is first produced as spermatogonia enter meiosis and differentiate into
119 spermatocytes[3,80]. To determine whether the MIWI-potentiating factor is differentially
120 expressed during spermatogenesis, we supplemented the target cleavage assay with
121 lysate prepared from FACS-purified germ cells: piRISC-potentiating activity was
122 greatest in lysate from secondary spermatocytes (Fig. 1c), a cell type in which
123 pachytene-piRNA-directed target cleavage is readily detected in vivo[47,50]. Moreover, the
124 potentiating activity was testis-specific: lysates from brain, liver, or kidney failed to
125 enhance piRNA-directed target cleavage by MIWI (Extended Data Fig. 3a). Finally, the
126 potentiating activity was specific for PIWI proteins and had no effect on the rate of target
127 cleavage by mouse AGO2 (Extended Data Fig. 3b).
128 Three lines of evidence suggest that the MIWI-potentiating activity contains one
129 or more structural Zn^{2+} ions. First, pre-treating testis lysate with the sulfhydryl alkylating
130 agent *N*-ethylmaleimide-inactivated the potentiating activity (Extended Data Fig. 3c),
131 indicating reduced cysteine residues, which often bind divalent metal cations[81], are
132 essential. Second, the MIWI-potentiating activity was unaltered by EGTA, which
133 chelates Ca^{2+}, but was irreversibly inactivated by EDTA, which chelates many divalent
134 metals, and by 1,10-Phenanthroline, which specifically chelates Zn^{2+} (Extended Data
135 Fig. 3d). Adding additional metal ions failed to rescue the activity (Extended Data Fig.
136 3e–g), suggesting that loss of Zn^{2+} irreversibly denatures the MIWI-potentiating factor, a
137 characteristic of zinc-finger proteins[82,83]. Third, the MIWI-potentiating activity bound
138 more tightly to an immobilized metal-affinity resin charged with Zn^{2+} than to resin
139 charged with Ni^{2+} (Extended Data Fig. 3h).
140 To identify the MIWI-potentiating activity, we developed a chromatographic
141 purification scheme using cation-exchange, hydrophobic-interaction, and size-exclusion
142 chromatography (Fig. 2a). Notably, the activity eluted from a Superdex 200 size-
143 exclusion column as a single ~17 kDa peak (Fig. 2b). Taken together, our data suggest

144 that the MIWI-potentiating activity corresponds to a small, testis-specific, Zn^{2+}-binding
145 protein abundantly expressed in meiotic and post-meiotic male germ cells (Fig. 2c).

146 **GTSF1 is necessary and sufficient for efficient target cleavage by MIWI**

147 Gametocyte Specific Factor 1 (GTSF1) is a conserved, 19,083 Da, tandem CHHC-type
148 zinc-finger protein essential for piRNA function and fertility in flies[7,8,10], silk moths[11],
149 worms[84], and mice[9,12]. Like *Miwi* and *Mili*, the mRNA abundance of *Gtsf1*, as well as its
150 paralogs *Gtsf1L* and *Gtsf2*, rises sharply as male germ cells enter meiosis I, peaks in
151 secondary spermatocytes, and then declines in round spermatids (Extended Data
152 Fig. 4a). Immobilized mouse GTSF1 captures all three PIWI proteins from testis lysate,
153 but GTSF1 and PIWI proteins co-expressed in HEK293T cells do not co-
154 immunoprecipitate, leading to the suggestion that the GTSF1-PIWI interaction requires
155 additional protein or RNA components[9]. MIWI2-bound piRNAs are lost in *Gtsf1*$^{-/-}$ male
156 mice[9], which are sterile[12] likely because loss of GTSF1 causes loss of MIWI2-directed
157 retrotransposon promoter methylation[4,85–87]. Although RNA cleavage by MIWI2 is not
158 required for piRNA biogenesis or function, production of MIWI2 piRISC requires MILI-
159 dependent piRNA amplification, a process that requires the endonuclease activity of
160 MILI[6]. In flies, piRNA-directed transcriptional silencing of transposons by Piwi, but not
161 Piwi piRISC-assembly, requires the GTSF1 ortholog Asterix[7,8].

162 To test whether the MIWI-potentiating activity in the testis lysate corresponded to
163 GTSF1, we used CRISPR-Cas9 to engineer a mouse strain with a 3XFLAG-HA epitope
164 tag inserted into the endogenous *Gtsf1* coding sequence (*Gtsf1*^{KI}; Fig. 2d). Because the
165 MIWI potentiating activity was greatest in secondary spermatocytes, we prepared lysate
166 from FACS-purified secondary spermatocytes from homozygous *Gtsf1*$^{KI/KI}$ male mice.
167 Incubation with anti-FLAG antibody coupled to paramagnetic beads depleted the lysate
168 of both epitope-tagged GTSF1 (Fig. 2e) and the MIWI-potentiating activity (Fig. 2f). The
169 activity was recovered from the beads by elution with 3XFLAG peptide (Fig. 2f). In

170 contrast, lysate from C57BL/6 secondary spermatocytes retained the MIWI-potentiating
171 factor after incubation with anti-FLAG antibody, and no activity was eluted from the
172 beads after incubation with FLAG peptide. Thus, GTSF1 is necessary to potentiate the
173 catalytic activity of MIWI.

174 GTSF1 is also sufficient to potentiate the catalytic activity of MIWI. Adding
175 purified recombinant GTSF1 (Extended Data Fig. 4b) to the MIWI-piRISC target
176 cleavage assay increased the rate of target cleavage by MIWI >100-fold (Fig. 3a, b).
177 The addition of 500 nM GTSF1 to MIWI piRISC (5 nM) caused the pre-steady-state rate
178 (k_{burst}) of target cleavage to increase from ~0.01 min^{-1} to 1.5 min^{-1} (Fig. 3b), a rate
179 similar to that of AGO2 (≥ 3 min^{-1})[74,75].

180 Our data suggest that GTSF1 does not promote product release or turnover: the
181 steady-state rate (k_{ss} <0.005 min^{-1}) of target cleavage under multiple-turnover
182 conditions (i.e., GTSF1 >> target RNA >> MIWI) was essentially unchanged from the
183 rate in the absence of GTSF1 (Fig. 1b and 3b). Perhaps MIWI remains bound to the
184 cleaved products, preventing it from catalyzing multiple rounds of target cleavage.
185 Supporting this idea, at incubation times >15 min, 3′-to-5′ exonucleases present in testis
186 lysate degrade the uncut, full-length target RNA, but the 5′ cleavage product remains
187 stable, consistent with its being protected by MIWI bound to its 3′ end (Fig. 1b). In vivo,
188 product release has been proposed to be facilitated in insects and mammals by the
189 RNA-stimulated ATPase Vasa (DDX4)[88–90] and the Vasa-like protein DDX43 (ref. 91).

190 **RNA-binding is essential for GTSF1 function**

191 Only three known eukaryotic proteins contain CHHC zinc-fingers: the spliceosomal
192 RNA-binding protein U11-48K, the TRM13 tRNA methyltransferase, and GTSF1 and its
193 paralogs[92]. In vitro, the first zinc-finger of GTSF1 directly binds RNA; RNA binding
194 requires four basic surface residues (R26, R29, K36 and K39)[93]. Potentiation of target
195 cleavage required RNA-binding by GTSF1: purified mutant $GTSF1^{R26A,R29A,K36A,K39A}$

196 (Extended Data Fig. 4b) did not detectably increase the rate of catalysis by either MIWI
197 or MILI (Fig. 3c–d and Extended Data Table 2), suggesting that GTSF1 must interact
198 with RNA to function.

199 **GTSF1 function requires PIWI-binding**

200 In flies and mice, GTSF1 interacts with PIWI proteins through conserved aromatic
201 residues in its carboxy-terminus[7,9] (Fig. 2c). Mutating these residues in the recombinant
202 protein—W98A, W107A, and W112A for mouse GTSF1—reduced the stimulatory effect
203 of mouse GTSF1 on MIWI in a concentration-dependent manner, supporting the idea
204 that GTSF1 binds MIWI directly when potentiating target cleavage (Fig. 3e and
205 Extended Data Table 2) and suggesting that W98, W107, and W112 define a conserved
206 surface of GTSF1-that interacts with PIWI proteins. We measured the pre-steady-state
207 catalytic rate (k_{burst}) as a function of GTSF1 concentration. Fitting the data to a
208 hyperbolic function revealed that GTSF1 binds MIWI piRISC >60-fold more tightly than
209 GTSF1^{W98A,W107A,W112A} (GTSF1: K_D = 8 nM, 95% confidence interval [C.I.] = 6 – 9 nM;
210 GTSF1^{W98A,W107A,W112A}: K_D = 500 nM, 95% C.I. = 100 – 900 nM) (Fig. 3e). Moreover,
211 PIWI-binding appears to be the sole defect in GTSF1^{W98A,W107A,W112A}, since the first-
212 order rates of target cleavage by MIWI at saturating concentrations of either wild-type or
213 mutant GTSF1 were essentially indistinguishable. (wild-type: 1.2 min^{-1}, 95% C.I. = 0.7 –
214 2 min^{-1}; mutant GTSF1: 0.8 min^{-1}, 95% C.I. = 0.6 – 0.9 min^{-1}).

215 **The PIWI-stimulatory function of GTSF1 orthologs is evolutionarily conserved**

216 GTSF1 orthologs can be found in many metazoan genomes[92], suggesting that GTSF1
217 may potentiate target cleavage by PIWI proteins in many animals. Supporting this idea,
218 purified recombinant Gtsf1 from the arthropod *Bombyx mori* (Extended Data Fig. 4c)
219 potentiated the catalytic activity of the *B. mori* PIWI protein, BmSiwi (Fig. 3f).

220 The structure and kinetics of EfPiwi from the sponge *Ephydatia fluviatilis* were
221 recently described[64]. Like MIWI, EfPiwi possesses inherently weak catalytic activity, a
222 feature common to all PIWI proteins examined to date. Although the *E. fluviatilis*
223 genome is yet to be sequenced, the closely related, fully sequenced genome of
224 *Ephydatia muelleri* contains a readily identifiable GTSF1 ortholog. Purified recombinant
225 EmGtsf1 (Extended Data Fig. 4c) stimulated the single-turnover catalytic rate of EfPiwi
226 piRISC ~6 fold (Fig. 3g). Sponges (Porifera) are the sister group to all other animals, the
227 Eumetazoa, having separated ~900 Mya. Thus, the last common ancestor of all animals
228 likely required GTSF1 to potentiate target cleavage by PIWI proteins.
229 The GTSF1 tandem zinc-finger domains are conserved across phyla, whereas
230 the GTSF1 carboxy-terminal sequence diverges substantially between mammals and
231 arthropods (Extended Data Table 1). For example, the sequences of the mouse and
232 rhesus macaque first and second zinc-fingers are 100% identical, whereas their C-
233 terminal domains share 88.5% identity. The first zinc-finger of mouse GTSF1 is 37.5%
234 and 45.8% identical to its fly and moth orthologs, but the mouse protein shares just 8%
235 and 8.3% identity with the C-terminal domains of the fly and moth proteins, respectively.
236 Consistent with the evolutionary divergence of their carboxy terminal domains, testis
237 lysate from rat or rhesus macaque enhanced target cleavage by mouse MIWI piRISC,
238 whereas lysate from *Drosophila melanogaster* or *Trichoplusia ni* ovaries, *T. ni* Hi5 cells,
239 or purified recombinant *Bombyx mori* BmGtsf1 did not (Extended Data Fig. 4c, 5a-b and
240 Supplementary Data Fig. 1d).

241 **GTSF1 paralogs can distinguish among PIWI proteins within a species**

242 Animal genomes often encode more than one GTSF protein[92] (Extended Data Fig. 5c).
243 For example, *D. melanogaster* has four GTSF paralogs. The *D. melanogaster* OSC and
244 OSS cell lines, which derive from somatic follicle cells that support oogenesis, express
245 Piwi but lack the PIWI paralogs Aub or Ago3. Piwi-mediated, piRNA-guided transposon

246 silencing in these cells requires Asterix, a GTSF1 ortholog[7,8,10]. In vivo, *asterix* mutants
247 phenocopy *piwi* mutants and are female sterile, even though Piwi is successfully loaded
248 with piRNAs and transits to the nucleus in the absence of Asterix[7,8]. Whether the other
249 fly GTSF paralogs have a function in vivo, perhaps as auxiliary factors for Aub or Ago3,
250 remains to be tested. Like fly Asterix, mouse GTSF1 is essential for piRNA function and
251 fertility. In mice, two GTSF1 paralogs, GTSF1L and GTSF2 are also expressed during
252 spermatogenesis and interact with PIWI proteins[94]. Unlike GTSF1, single and double
253 *Gtsf1l* and *Gtsf2* knockout males are fertile[94]. Genes encoding the *Gtsf* paralogs are
254 syntenic in mammals, while *Gtsf2* is lost in primates (Extended Data Fig. 6).
255 The C-terminal domains of GTSF orthologs are more similar among closely
256 related species than among GTSF paralogs within the same species (Extended Data
257 Fig. 4d and 5c and Extended Data Table 1), further supporting the view that this domain
258 has evolved to bind specific PIWI proteins. Consistent with this idea, mouse GTSF1,
259 GTSF1L, and GTSF2 differ in their ability to potentiate target cleavage by MIWI and
260 MILI. While GTSF1 accelerated target cleavage by both MIWI and MILI, purified,
261 recombinant GTSF1L and GTSF2 (Extended Data Fig. 4b) efficiently potentiated target
262 cleavage by MIWI but not MILI (Fig. 3c–d and Extended Data Table 2). GTSF2 was
263 unable to detectably increase the rate of target cleavage by MILI. Although GTSF1L had
264 a modest effect on the rate of target cleavage by MILI piRISC, this enhancement was
265 less than one-sixteenth that provided by GTSF1 and half that of the PIWI-binding
266 mutant GTSF1^{W98A,W107A,W112A} (Fig. 3c–d and Extended Data Table 2). We conclude that
267 GTSF1L and GTSF2 are specialized to potentiate target cleavage by MIWI piRISC.
268 Silk moth BmGtsf1-like is more similar to mouse GTSF1 than to BmGtsf1
269 (Extended Data Fig. 5c and Extended Data Table 1); BmGtsf1-like more than doubled
270 the amount of target cleavage by MIWI piRISC, but had no detectable effect on MILI
271 (Extended Data Fig. 5b), further supporting the idea that the GTSF carboxy terminal
272 domain determines the affinity of the protein for specific PIWI proteins. In vivo, BmGtsf1

273 interacts with Siwi and is required for transposon silencing and sex determination[11].
274 BmGtsf1 increased the rate of target cleavage by affinity-purified BmSiwi but not that of
275 the other silk moth PIWI protein, BmAgo3 (Fig. 3f).

276 **Target cleavage by MIWI or MILI requires extensive pairing between piRNA and**
277 **target**

278 piRNA:target RNA complementarity from g2–g22 (i.e., 21 base-pairs) is required for
279 efficient target cleavage directed by endogenous piRNAs loaded into MIWI piRISC
280 immunoprecipitated from adult mouse testis[5]. However, GTSF1 does not detectably co-
281 immunoprecipitate with MIWI piRISC from mouse testis[5,77]. A requirement for 21 bp
282 complementarity between target and guide is unprecedented among Argonaute
283 proteins: fly Ago2 slices a target with as few as 11 contiguous base pairs[95]; mammalian
284 AGO2 requires only 11 contiguous base pairs for detectable cleavage[96]; and the
285 eubacterial DNA-guided DNA endonuclease TtAgo requires as few as 14 (ref. 97).
286 Affinity-purified MIWI, programmed with either of two different synthetic, 30 nt
287 piRNAs, readily cleaved target RNA complementary to guide nucleotides g2–21 but not
288 a target complementary to g2–g16 (Extended Data Fig. 7a, b). Under multiple-turnover
289 conditions with saturating amounts of purified, recombinant GTSF1 ([GTSF1] >> [target]
290 > [piRISC]), MIWI readily cleaved a target RNA with 19 nucleotides (g2–g20)
291 complementary to its synthetic piRNA guide (Fig. 4a, top). The lower background of
292 single-turnover experiments ([GTSF1] >> [piRISC] > [target]) allowed longer incubation
293 times. Using these conditions, we could detect GTSF1-stimulated cleavage of a target
294 RNA with as few as 15 complementary nucleotides (g2–g16; Fig. 4a, bottom). We note
295 that the pachytene stage of meiosis in mouse spermatogenesis lasts about 175 h, and
296 the pachytene piRNA pathway components are expressed until at least the round
297 spermatid stage, a time interval spanning >400 h.

298 **Guide length limits the rate of target cleavage by MIWI**

299 In vivo, piRNAs are trimmed to a length characteristic of the PIWI protein in which they
300 reside: ~30 nt for MIWI and ~26–27 nt for MILI[41,42,87]. An attractive hypothesis is that
301 these piRNA lengths are optimal for target cleavage catalyzed by the specific PIWI
302 protein. Our data suggest a more complex relationship between piRNA length, PIWI
303 protein identity, and target complementarity. In the presence of GTSF1, MIWI loaded
304 with a 30 nt piRNA and MILI loaded with a 26 nt piRNA readily cleaved a fully
305 complementary target RNA in a 60 min reaction (Fig. 4b). By contrast, neither piRISC
306 cleaved a target complementary to piRNA positions g2–g16 (Fig. 4a, b and Extended
307 Data Fig. 7a, b). Similarly, MIWI loaded with a 26 nt or MILI loaded with a 26 or 21 nt
308 guide produced little cleavage for a target complementary to positions g2–g16 although
309 both guide lengths supported cleavage of a fully complementary target (Fig. 4b). In fact,
310 for MIWI, a 21mer was more active than a 30 nt guide (Fig. 4b). Without GTSF1, MIWI
311 or MILI loaded with any of these guide lengths produced little cleaved target in 60 min
312 (Fig. 4b). Remarkably, MIWI or MILI loaded with a 16 nt guide RNA, a piRNA length not
313 present in vivo, readily cleaved the RNA target (Fig. 4b).

314 GTSF1 accelerates the rate of pre-steady state target cleavage (k_{burst}) by MIWI
315 and MILI but has little effect on the nearly slow rate of steady-state cleavage (k_{ss}),
316 suggesting that piRISC remains bound to its cleavage by-products (Fig. 3e). We
317 incubated GTSF1 and MIWI—loaded with a 30, 26, 21, or 16 nt guide—with a target
318 RNA fully complementary to each guide and measured k_{burst} and k_{ss} (Fig. 4c). As the
319 guide length decreased, k_{burst} decreased and k_{ss} increased. Compared to its native 30 nt
320 guide length, the 16 nt guide increased k_{ss} ~9-fold and decreased k_{burst} >50-fold. These
321 data suggest that as the guide was shortened, the cleaved products were released
322 more rapidly.

323 We estimated the binding affinity (ΔG) of the piRNA for its target using the
324 nearest-neighbor rules for base pairing at 37°C. As the strength of the piRNA base
325 pairing to the target increased, the rate of pre-steady-state cleavage increased (Fig. 5a),
326 consistent with base pairing serving to extract the 3′ end of the piRNA from the PAZ
327 domain, facilitating the transition to a more catalytically competent piRISC
328 conformation[98]. Conversely, decreased base-pairing strength increased the steady-state
329 rate of target cleavage, supporting the view that for biologically relevant piRNA lengths,
330 product release is the rate-determining step for MIWI-catalyzed target cleavage
331 (Fig. 5a).

332 **Discussion**

333 In nearly all animals, both piRNA biogenesis and piRNA function require the PIWI
334 endoribonuclease activity. Yet purified mouse MIWI and MILI are intrinsically slow to
335 cleave complementary target RNAs. Our data show that unlike AGO proteins, PIWI
336 proteins require an auxiliary factor, GTSF1, to efficiently cleave their RNA targets.
337 The ability of GTSF1 to potentiate PIWI-catalyzed target cleavage provides a
338 biochemical explanation for the genetic requirement for this small zinc-finger protein in
339 the piRNA pathway. GTSF1 function requires that it bind both RNA and the PIWI
340 protein, and differences among carboxy-terminal domains restrict individual GTSF1
341 paralogs to specific PIWI proteins.
342 We propose a testable kinetic scheme for target cleavage by MIWI (Fig. 5b, c
343 and Extended Data Fig. 8) that incorporates the requirement for GTSF1 and the
344 observation that a 16 nt guide changes the rate-determining step of target cleavage
345 catalyzed by MIWI. As originally proposed for fly Ago2 (ref. 98), piRISC bound to a
346 target is presumed to exist in two-states: one in which the piRNA 3′ end is secured in
347 the PAZ domain and a competing, pre-catalytic conformation in which the piRNA is fully
348 base paired to its target. Structures of Piwi-A from the freshwater sponge *Ephydatia*

349 *fluviatilis* show that extensive pairing between a piRNA and its target induces a
350 catalytically competent conformation in which the PAZ domain is rotated away from the
351 piRNA 3′ end[64]. We propose that the PAZ-bound state is more favorable for guides
352 bound to PIWI proteins than to AGOs, requiring a high degree of complementarity
353 between guide and target to extract the piRNA 3′ terminus from the PAZ domain. Our
354 data also suggest that a slow rate of product release after cleavage results directly from
355 the extensive piRNA-target RNA complementarity required for this conformational
356 change. A 16 nt piRNA allows MIWI more easily to assume a pre-catalytic state,
357 perhaps because the 3′ end of the short guide cannot reach the PAZ domain. Notably,
358 single-stranded siRNA guides as short as 14 nt allow mammalian AGO3, initially
359 believed to have lost its endonuclease activity, to efficiently cleave RNA targets[99]. In
360 golden hamsters, piRNAs bound to PIWIL1 are initially ~29 nt long, but a shorter
361 population of ~23 nt piRNAs appears at metaphase II and predominates in two-cell
362 embryos[100]. piRNAs bound to PIWIL3, a female-specific PIWI protein absent from mice,
363 are ~19 nt long in hamster and ~20 nt in human oocytes[100,101]. We speculate that these
364 short piRNAs allow PIWIL1 and PIWIL3 to function as multiple-turnover endonucleases.
365 Our model proposes that GTSF1 recognizes the pre-catalytic piRISC state,
366 facilitating a second conformational change in PIWI proteins that may occur
367 spontaneously in AGO clade proteins. GTSF1 binding likely stabilizes the catalytically
368 active conformation, facilitating target cleavage, but has no detectable effect on
369 subsequent release of the cleaved products. Slow product release may be a general
370 property of PIWI proteins: purified *Ephydatia fluviatilis* (freshwater sponge) PIWI
371 similarly catalyzes only a single round of target cleavage[64].
372

373 **References**

374 1. Ozata, D. M., Gainetdinov, I., Zoch, A., O'Carroll, D. & Zamore, P. D. PIWI-
375 interacting RNAs: small RNAs with big functions. *Nat. Rev. Genet.* **20**, 89-108
376 (2018).

377 2. Kuramochi-Miyagawa, S. et al. *Mili*, a mammalian member of *piwi* family gene, is
378 essential for spermatogenesis. *Development* **131**, 839-849 (2004).

379 3. Deng, W. & Lin, H. *miwi*, a murine homolog of *piwi*, encodes a cytoplasmic protein
380 essential for spermatogenesis. *Dev. Cell* **2**, 819-830 (2002).

381 4. Carmell, M. A. et al. MIWI2 is essential for spermatogenesis and repression of
382 transposons in the mouse male germline. *Dev. Cell* **12**, 503-514 (2007).

383 5. Reuter, M. et al. Miwi catalysis is required for piRNA amplification-independent
384 LINE1 transposon silencing. *Nature* **480**, 264-267 (2011).

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

387 7. Dönertas, D., Sienski, G. & Brennecke, J. *Drosophila* Gtsf1 is an essential
388 component of the Piwi-mediated transcriptional silencing complex. *Genes Dev.* **27**,
389 1693-1705 (2013).

390 8. Ohtani, H. et al. DmGTSF1 is necessary for Piwi-piRISC-mediated transcriptional
391 transposon silencing in the *Drosophila* ovary. *Genes Dev.* **27**, 1656-1661 (2013).

392 9. Yoshimura, T. et al. Mouse GTSF1 is an essential factor for secondary piRNA
393 biogenesis. *EMBO Rep.* **19**, e42054 (2018).

394 10. Muerdter, F. et al. A genome-wide RNAi screen draws a genetic framework for
395 transposon control and primary piRNA biogenesis in *Drosophila*. *Mol. Cell.* **50**,
396 736-748 (2013).

397 11. Chen, K. et al. Gtsf1 is essential for proper female sex determination and
398 transposon silencing in the silkworm, *Bombyx mori*. *PLoS Genet.* **16**, e1009194
399 (2020).

400 12. Yoshimura, T. et al. *Gtsf1/Cue110*, a gene encoding a protein with two copies of a
401 CHHC Zn-finger motif, is involved in spermatogenesis and retrotransposon
402 suppression in murine testes. *Dev. Biol.* **335**, 216-227 (2009).

403 13. Mohn, F., Handler, D. & Brennecke, J. Noncoding RNA. piRNA-guided slicing
404 specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. *Science*
405 **348**, 812-817 (2015).

406 14. Han, B. W., Wang, W., Li, C., Weng, Z. & Zamore, P. D. Noncoding RNA. piRNA-
407 guided transposon cleavage initiates Zucchini-dependent, phased piRNA
408 production. *Science* **348**, 817-821 (2015).

409 15. Homolka, D. et al. PIWI Slicing and RNA Elements in Precursors Instruct
410 Directional Primary piRNA Biogenesis. *Cell Rep* **12**, 418-428 (2015).

411 16. Gainetdinov, I., Colpan, C., Arif, A., Cecchini, K. & Zamore, P. D. A Single
412 Mechanism of Biogenesis, Initiated and Directed by PIWI Proteins, Explains piRNA
413 Production in Most Animals. *Mol. Cell.* **71**, 775-790.e5 (2018).

414 17. Ipsaro, J. J., Haase, A. D., Knott, S. R., Joshua-Tor, L. & Hannon, G. J. The
415 structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis.
416 *Nature* **491**, 279-283 (2012).

417 18. Ding, D. et al. PNLDC1 is essential for piRNA 3′ end trimming and transposon
418 silencing during spermatogenesis in mice. *Nat Commun* **8**, 819 (2017).

419 19. Zhang, Y. et al. An essential role for PNLDC1 in piRNA 3′ end trimming and male
420 fertility in mice. *Cell Res* **27**, 1392-1396 (2017).

421 20. Nishimura, T. et al. PNLDC1, mouse pre-piRNA Trimmer, is required for meiotic
422 and post-meiotic male germ cell development. *EMBO Rep.* **19**, (2018).

423 21. Wang, H. et al. Antagonistic roles of Nibbler and Hen1 in modulating piRNA 3'
424 ends in *Drosophila*. *Development* **143**, 530-539 (2016).

425 22. Feltzin, V. L. et al. The exonuclease Nibbler regulates age-associated traits and
426 modulates piRNA length in *Drosophila*. *Aging Cell* **14**, 443-452 (2015).

427 23. Izumi, N., Shoji, K., Suzuki, Y., Katsuma, S. & Tomari, Y. Zucchini consensus
428 motifs determine the mechanism of pre-piRNA production. *Nature* **578**, 311-316
429 (2020).

430 24. Izumi, N. et al. Identification and Functional Analysis of the Pre-piRNA 3' Trimmer
431 in Silkworms. *Cell* **164**, 962-973 (2016).

432 25. Haase, A. D. et al. Probing the initiation and effector phases of the somatic piRNA
433 pathway in *Drosophila*. *Genes Dev.* **24**, 2499-2504 (2010).

434 26. Gunawardane, L. S. et al. A Slicer-Mediated Mechanism for Repeat-Associated
435 siRNA 5′ End Formation *in Drosophila*. *Science* **315**, 1587-1590 (2007).

436 27. Brennecke, J. et al. Discrete Small RNA-Generating Loci as Master Regulators of
437 Transposon Activity in *Drosophila*. *Cell* **128**, 1089-1103 (2007).

438 28. Lewis, S. H. et al. Pan-arthropod analysis reveals somatic piRNAs as an ancestral
439 defence against transposable elements. *Nat Ecol Evol* **2**, 174-181 (2018).

440 29. Fu, Y. et al. The genome of the Hi5 germ cell line from *Trichoplusia ni*, an
441 agricultural pest and novel model for small RNA biology. *eLife* **7**, e31628 (2018).

442 30. Jehn, J. et al. PIWI genes and piRNAs are ubiquitously expressed in mollusks and
443 show patterns of lineage-specific adaptation. *Commun Biol* **1**, 137 (2018).

444 31. Lim, R. S., Anand, A., Nishimiya-Fujisawa, C., Kobayashi, S. & Kai, T. Analysis of
445 Hydra PIWI proteins and piRNAs uncover early evolutionary origins of the piRNA
446 pathway. *Dev. Biol.* **386**, 237-251 (2014).

447 32. Juliano, C. E. et al. PIWI proteins and PIWI-interacting RNAs function in Hydra
448 somatic stem cells. *Proc. Natl. Acad. Sci. U S A* **111**, 337-342 (2014).

449 33. Praher, D. et al. Characterization of the piRNA pathway during development of the
450 sea anemone *Nematostella vectensis*. *RNA Biol* **14**, 1727-1741 (2017).

451 34. Grimson, A. et al. Early origins and evolution of microRNAs and Piwi-interacting
452 RNAs in animals. *Nature* **455**, 1193-1197 (2008).

453 35. Kerner, P., Degnan, S. M., Marchand, L., Degnan, B. M. & Vervoort, M. Evolution
454 of RNA-binding proteins in animals: insights from genome-wide analysis in the
455 sponge *Amphimedon queenslandica*. *Mol. Biol. Evol.* **28**, 2289-2303 (2011).

456 36. Murchison, E. P. et al. Conservation of small RNA pathways in platypus. *Genome*
457 *Res.* **18**, 995-1004 (2008).

458 37. Houwing, S. et al. A role for Piwi and piRNAs in germ cell maintenance and
459 transposon silencing in Zebrafish. *Cell* **129**, 69-82 (2007).

460 38. Toombs, J. A. et al. Xenopus Piwi proteins interact with a broad proportion of the
461 oocyte transcriptome. *RNA* **23**, 504-520 (2017).

462 39. Sun, Y. H. et al. Domestic chickens activate a piRNA defense against avian
463 leukosis virus. *eLife* **6**, e24695 (2017).

464 40. Grivna, S. T., Beyret, E., Wang, Z. & Lin, H. A novel class of small RNAs in mouse
465 spermatogenic cells. *Genes Dev.* **20**, 1709-1714 (2006).

466 41. Aravin, A. et al. A novel class of small RNAs bind to MILI protein in mouse testes.
467 *Nature* **442**, 203-207 (2006).

468 42. Girard, A., Sachidanandam, R., Hannon, G. J. & Carmell, M. A. A germline-specific
469 class of small RNAs binds mammalian Piwi proteins. *Nature* **442**, 199-202 (2006).

470 43. Lim, S. L. et al. Conservation and expression of PIWI-interacting RNA pathway
471 genes in male and female adult gonad of amniotes. *Biol. Reprod.* **89**, 136 (2013).

472 44. Lau, N. C. et al. Characterization of the piRNA complex from rat testes. *Science*
473 **313**, 363-367 (2006).

474 45. Özata, D. M. et al. Evolutionarily conserved pachytene piRNA loci are highly
475 divergent among modern humans. *Nat Ecol Evol* **4**, 156-168 (2020).

476 46. Li, X. Z. et al. An Ancient Transcription Factor Initiates the Burst of piRNA
477 Production during Early Meiosis in Mouse Testes. *Mol. Cell.* **50**, 67-81 (2013).

478 47. Wu, P. H. et al. The evolutionarily conserved piRNA-producing locus *pi6* is
479 required for male mouse fertility. *Nat. Genet.* **52**, 728-739 (2020).

480 48. Choi, H., Wang, Z. & Dean, J. Sperm acrosome overgrowth and infertility in mice
481 lacking chromosome 18 pachytene piRNA. (2020).

482 49. Gou, L. T. et al. Pachytene piRNAs instruct massive mRNA elimination during late
483 spermiogenesis. *Cell Res* 1-21 (2014).

484 50. Goh, W. S. et al. piRNA-directed cleavage of meiotic transcripts regulates
485 spermatogenesis. *Genes Dev.* **29**, 1032-1044 (2015).

486 51. Vourekas, A. et al. Mili and Miwi target RNA repertoire reveals piRNA biogenesis
487 and function of Miwi in spermiogenesis. *Nat. Struct. Mol. Biol.* **19**, 773-781 (2012).

488 52. Zhang, P. et al. MIWI and piRNA-mediated cleavage of messenger RNAs in
489 mouse testes. *Cell Res* **25**, 193-207 (2015).

490 53. Wang, Y., Sheng, G., Juranek, S., Tuschl, T. & Patel, D. J. Structure of the guide-
491 strand-containing argonaute silencing complex. *Nature* **456**, 209-213 (2008).

492 54. Schirle, N. T. & MacRae, I. J. The crystal structure of human Argonaute2. *Science*
493 **336**, 1037-1040 (2012).

494 55. Matsumoto, N. et al. Crystal Structure of Silkworm PIWI-Clade Argonaute Siwi
495 Bound to piRNA. *Cell* **167**, 484-497.e9 (2016).

496 56. Parker, J. S., Roe, S. M. & Barford, D. Crystal structure of a PIWI protein suggests
497 mechanisms for siRNA recognition and slicer activity. *EMBO J.* **23**, 4727-4737
498 (2004).

499 57. Ma, J. B. et al. Structural basis for 5′-end-specific recognition of guide RNA by th*e*
500 *A. fulgidus* Piwi protein. *Nature* **434**, 663-666 (2005).

501 58. Yuan, Y. R. et al. Crystal structure of *A. aeolicus* Argonaute, a site-specific DNA-
502 guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage.
503 *Mol. Cell.* **19**, 405-419 (2005).

504 59. Ma, J. B. et al. Structural basis for 5′-end-specific recognition of guide RNA by the
505 *A. fulgidus* Piwi protein. *Nature* **434**, 666-670 (2005).

506 60. Le Thomas, A., Marinov, G. & Aravin, A. A Transgenerational Process Defines
507 piRNA Biogenesis in *Drosophila virilis*. *Cell Reports* **8**, 1617-1623 (2014).

508 61. Yamaguchi, S. et al. Crystal structure of Drosophila Piwi. *Nat Commun* **11**, 858
509 (2020).

510 62. Kawaoka, S., Izumi, N., Katsuma, S. & Tomari, Y. 3′ end formation of PIWI-
511 interacting RNAs in vitro. *Mol. Cell.* **43**, 1015-1022 (2011).

512 63. Cora, E. et al. The MID-PIWI module of Piwi proteins specifies nucleotide- and
513 strand-biases of piRNAs. *RNA* **20**, 773-81 (2014).

514 64. Anzelon, T. A. et al. Structural basis for piRNA targeting. *Nature* **597**, 285-289
515 (2021).

516 65. Kirino, Y. & Mourelatos, Z. The mouse homolog of HEN1 is a potential methylase
517 for Piwi-interacting RNAs. *RNA* **13**, 1397-1401 (2007).

518 66. Kirino, Y. & Mourelatos, Z. 2′-*O*-methyl modification in mouse piRNAs and its
519 methylase. *Nucleic Acids Symp Ser (Oxf)* 417-418 (2007).

520 67. Lim, S. L. et al. HENMT1 and piRNA Stability Are Required for Adult Male Germ
521 Cell Transposon Repression and to Define the Spermatogenic Program in the
522 Mouse. *PLoS Genet.* **11**, e1005620 (2015).

523 68. Kamminga, L. M. et al. Hen1 is required for oocyte development and piRNA
524 stability in zebrafish. *EMBO J.* **29**, 3688-3700 (2010).

525 69. Montgomery, T. A. et al. PIWI associated siRNAs and piRNAs specifically require
526 the *Caenorhabditis elegans* HEN1 ortholog henn-1. *PLoS Genet.* **8**, e1002616
527 (2012).

528 70. Billi, A. C. et al. The *Caenorhabditis elegans* HEN1 ortholog, HENN-1, methylates
529 and stabilizes select subclasses of germline small RNAs. *PLoS Genet.* **8**,
530 e1002617 (2012).

531 71. Kamminga, L. M. et al. Differential impact of the HEN1 homolog HENN-1 on 21U
532 and 26G RNAs in the germline of *Caenorhabditis elegans*. *PLoS Genet.* **8**,
533 e1002702 (2012).

534 72. Pelisson, A., Sarot, E., Payen-Groschene, G. & Bucheton, A. A novel repeat-
535 associated small interfering RNA-mediated silencing pathway downregulates
536 complementary sense *gypsy* transcripts in somatic cells of the *Drosophila* ovary. *J.*
537 *Virol.* **81**, 1951-1960 (2007).

538 73. Ameres, S. L. et al. Target RNA-directed trimming and tailing of small silencing
539 RNAs. *Science* **328**, 1534-1539 (2010).

540 74. Salomon, W. E., Jolly, S. M., Moore, M. J., Zamore, P. D. & Serebrov, V. Single-
541 Molecule Imaging Reveals that Argonaute Reshapes the Binding Properties of Its
542 Nucleic Acid Guides. *Cell* **162**, 84-95 (2015).

543 75. Becker, W. R. et al. High-Throughput Analysis Reveals Rules for Target RNA
544 Binding and Cleavage by AGO2. *Mol. Cell.* **75**, 741-755.e11 (2019).

545 76. Hutvagner, G. & Zamore, P. D. A microRNA in a multiple-turnover RNAi enzyme
546 complex. *Science* **297**, 2056-2060 (2002).

547 77. Vagin, V. V. et al. Proteomic analysis of murine Piwi proteins reveals a role for
548 arginine methylation in specifying interaction with Tudor family members. *Genes*
549 *Dev.* **23**, 1749-1762 (2009).

550 78. Siomi, M. C., Mannen, T. & Siomi, H. How does the royal family of Tudor rule the
551 PIWI-interacting RNA pathway? *Genes Dev.* **24**, 636-646 (2010).

552 79. Bedford, M. T. & Clarke, S. G. Protein arginine methylation in mammals: who,
553 what, and why. *Mol. Cell.* **33**, 1-13 (2009).

554 80. Kuramochi-Miyagawa, S. et al. Two mouse *piwi*-related genes: *miwi* and *mili*. *Mech*
555 *Dev* **108**, 121-133 (2001).

556 81. Michael, S. F., Kilfoil, V. J., Schmidt, M. H., Amann, B. T. & Berg, J. M. Metal
557 binding and folding properties of a minimalist Cys2His2 zinc finger peptide. *Proc.*
558 *Natl. Acad. Sci. U S A* **89**, 4796-4800 (1992).

559 82. Matt, T., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. The CBP/p300
560 TAZ1 domain in its native state is not a binding partner of MDM2. *Biochem. J.* **381**,
561 685-691 (2004).

562 83. Nyborg, J. K. & Peersen, O. B. That zincing feeling: the effects of EDTA on the
563 behaviour of zinc-binding transcriptional regulators. *Biochem. J.* **381**, e3-4 (2004).

564 84. Almeida, M. V. et al. GTSF-1 is required for formation of a functional RNA-
565 dependent RNA Polymerase complex in *Caenorhabditis elegans. EMBO J.* **37**,
566 e99325 (2018).

567 85. Kuramochi-Miyagawa, S. et al. DNA methylation of retrotransposon genes is
568 regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes*
569 *Dev.* **22**, 908-917 (2008).

570 86. Kojima-Kita, K. et al. MIWI2 as an Effector of DNA Methylation and Gene Silencing
571 in Embryonic Male Germ Cells. *Cell Rep* **16**, 2819-2828 (2016).

572 87. Aravin, A. A. et al. A piRNA pathway primed by individual transposons is linked to
573 de novo DNA methylation in mice. *Mol. Cell.* **31**, 785-799 (2008).

574 88. Wenda, J. M. et al. Distinct Roles of RNA Helicases MVH and TDRD9 in PIWI
575 Slicing-Triggered Mammalian piRNA Biogenesis and Function. *Dev. Cell* **41**, 623-
576 637.e9 (2017).

577 89. Xiol, J. et al. RNA Clamping by Vasa Assembles a piRNA Amplifier Complex on
578 Transposon Transcripts. *Cell* (2014).

579 90. Nishida, K. M. et al. Respective functions of two distinct Siwi complexes
580 assembled during PIWI-interacting RNA biogenesis in *Bombyx* germ cells. *Cell*
581 *Rep* **10**, 193-203 (2015).

582 91. Murakami, R., Sumiyoshi, T., Negishi, L. & Siomi, M. C. DEAD-box polypeptide 43
583 facilitates piRNA amplification by actively liberating RNA from Ago3-piRISC. *EMBO*
584 *Rep.* **22**, e51313 (2021).

585 92. Andreeva, A. & Tidow, H. A novel CHHC Zn-finger domain found in spliceosomal
586 proteins and tRNA modifying enzymes. *Bioinformatics* **24**, 2277-2280 (2008).

587 93. Ipsaro, J. J., O'Brien, P. A., Bhattacharya, S., Palmer III, A. G. & Joshua-Tor, L.
588 Asterix/Gtsf1 links tRNAs and piRNA silencing of retrotransposons. *Cell Reports*
589 **34**, 108914 (2021).

590 94. Takemoto, N., Yoshimura, T., Miyazaki, S., Tashiro, F. & Miyazaki, J. *Gtsf1l and*
591 Gtsf2 Are Specifically Expressed in Gonocytes and Spermatids but Are Not
592 Essential for Spermatogenesis. *PLoS One* **11**, e0150390 (2016).

593 95. Haley, B. & Zamore, P. D. Kinetic analysis of the RNAi enzyme complex. *Nat.*
594 *Struct. Mol. Biol.* **11**, 599-606 (2004).

595 96. Jackson, A. L. et al. Expression profiling reveals off-target gene regulation by
596 RNAi. *Nat. Biotechnol.* **21**, 635-637 (2003).

597 97. Wang, Y. et al. Nucleation, propagation and cleavage of target RNAs in Ago
598 silencing complexes. *Nature* **461**, 754-761 (2009).

599 98. Tomari, Y. & Zamore, P. D. Perspective: machines for RNAi. *Genes Dev.* **19**, 517-
600 529 (2005).

601 99. Park, M. S., Sim, G., Kehling, A. C. & Nakanishi, K. Human Argonaute2 and
602 Argonaute3 are catalytically activated by different lengths of guide RNA.
603 *Proceedings of the National Academy of Sciences* **117**, 28576-28578 (2020).

604 100. Ishino, K. et al. Hamster PIWI proteins bind to piRNAs with stage-specific size
605 variations during oocyte maturation. *bioRxiv* 2020.12.01.407411; doi:
606 https://doi.org/10.1101/2020.12.01.407411 (2020).

607 101. Yang, Q. et al. Single-cell CAS-seq reveals a class of short PIWI-interacting RNAs
608 in human oocytes. *Nature Comm.* **10**, 1-15 (2019).

609

610 **Methods**

611 **Plasmids and cell lines**

612 Supplementary Data Table 1 provides the sequences of all oligonucleotides used. To
613 create pScalps_Puro_EGFP, an IRES-driven EGFP was inserted downstream of the
614 multiple cloning site and upstream of the puromycin coding sequence of the lentivirus
615 transfer vector pScalps_Puro. MIWI cDNA was obtained from Mammalian Gene
616 Collection (https://genecollections.nci.nih.gov/MGC/). MILI cDNA was amplified by RT-
617 PCR from mouse testis total RNA. Gibson assembly and restriction cloning were used
618 to clone the MILI and MIWI coding sequences into pScalps_Puro_EGFP, fusing them
619 in-frame with N-terminal 3XFLAG and SNAP tags. Lentivirus transfer vectors were
620 packaged by co-transfection with psPAX2 and pMD2.G (4:3:1) using TransIT-2020
621 (Mirus Bio) in HEK293T cells. Supernatant containing lentivirus was used to transduce
622 HEK293T cells in the presence of 16 ug/ml polybrene (Sigma) to obtain stable PIWI-
623 expressing cell lines. Three sequential transductions were performed to maximize
624 recombinant protein production. The transduced cells were selected in the presence of
625 2 ug/ml Puromycin for two weeks, then the cells expressing the 5–10% highest EGFP
626 fluorescence were selected by FACS (UMASS Medical School Flow Cytometry Core).
627 Selected cells stably expressing the recombinant PIWI proteins were expanded,
628 harvested, and cell pellets flash-frozen and stored at −80°C.

629 Mouse GTSF1 cDNA was synthesized at Twist Biosciences and cloned into
630 pCold-GST (Takara Bio) bacterial expression vector by restriction cloning. GTSF1
631 mutants, GTSF1L, and GTSF2- expressing pCold-GST vectors were synthesized at
632 Twist Biosciences. pIZ-FLAG6His-Siwi was described previously[102]. FLAG-tagged MILI
633 and MIWI-expressing vectors were the kind gift of Shinpei Kawaoka (Kyoto University,
634 Kyoto, Japan). MmGtsf1 cDNA was amplified by RT-PCR from mouse spermatogonial
635 stem cell total RNA[103]. BmGtsf1 and BmGtsf1-like cDNAs were amplified by RT-PCR
636 from BmN4 cell total RNA. The amplified cDNA fragment and a DNA fragment coding

637 V5SBP were cloned into pcDNA5/FRT/TO vector (Thermo Fisher Scientific) by In-fusion
638 cloning (Takara). The *Ephydatia muelleri* Gtsf1 coding sequence was obtained from [104]
639 and cloned into bacterial expression vector pSV272 (IJM).

640 **Mice**

641 Generation of 3XFLAG-HA-tagged GTSF1 mice (*Gtsf1*$^{KI/KI}$) was performed at Cyagen.
642 The coding sequence for the tags was inserted into the endogenous locus by CRISPR-
643 Cas9. Briefly, fertilized mouse embryos were injected with the sgRNA targeting the
644 sequence GTCTTCCATGCTGATGGCAAAGG (PAM), a 3XFLAG-HA-tag cassette HDR
645 donor, and Cas9 mRNA. Supplementary Data Table 1 provides the sequences of the
646 HDR donor and oligonucleotide primers used for genotyping. Founder F0 animals were
647 genotyped and bred to generate F1 animals carrying the germline-transmitted knock-in
648 allele. Mice were maintained and used according to the guidelines of the Institutional
649 Animal Care and Use Committee of the University of Massachusetts Chan Medical
650 School (A201900331). C57BL/6J mice (RRID: IMSR_JAX:000664) were used as wild-
651 type control, where indicated.

652 **Recombinant protein purification**

653 *PIWI proteins.* PIWI-expressing stable cells were harvested by centrifugation and stored
654 at −80°C until lysed in 10 ml lysis buffer (30 mM HEPES-KOH, pH 7.5, 100 mM
655 potassium acetate, 3.5 mM magnesium acetate, 1 mM DTT, 0.1% v/v Triton X-100,
656 20% v/v glycerol, and 1× protease inhibitor cocktail [1 mM 4-(2-
657 aminoethyl)benzenesulfonyl fluoride hydrochloride [Sigma; A8456], 0.3 μM Aprotinin, 40
658 μM betanin hydrochloride, 10 μM. E-64 [Sigma; E3132], 10 μM leupeptin hemisulfate])
659 per g frozen cells. Cell lysis was monitored by staining with trypan blue. Crude
660 cytoplasmic lysate was clarified at 20,000 (S20) or 100,000 × *g* (S100) for 30 mins,
661 aliquoted, flash-frozen, and stored at −80°C. To capture PIWI proteins, clarified lysate

662 was incubated with 20 µl Anti-FLAG M2 paramagnetic beads (Sigma) per ml of lysate
663 for 2 h to overnight rotating at 4°C. Beads were washed five times with high salt wash
664 buffer (30 mM HEPES-KOH, pH 7.5, 2 M potassium acetate, 3.5 mM magnesium
665 acetate, 1 mM DTT), twice with low salt buffer (high salt wash buffer except containing
666 100 mM potassium acetate and 0.01% v/v Triton X-100). To assemble RISC, beads
667 were resuspended in low salt buffer containing 100 nM synthetic guide piRNA and
668 incubated with rotation at 37°C or room temperature for 30 min. MIWI RISC or unloaded
669 apo-MIWI was eluted from the beads with 200 ng/ul 3XFLAG peptide in lysis buffer
670 without 0.1% v/v Triton X-100 or protease inhibitors for 2 h at 4°C. Eluate containing
671 PIWI piRISC was aliquoted and stored at −80°C. EfPiwi piRISC purification has been
672 described[64]. AGO2 siRISC was purified using oligo-affinity purification as described[105].

673 *GTSF1, GTSF1 mutants, and GTSF1 homologs*. pCold-GST GTSF-expression vectors
674 were transformed into Rosetta-Gami 2 competent cells (Novagen). Cells were grown in
675 the presence of 1 µM $ZnSO_4$ at 37°C until OD_{600} ~0.6–0.8, then chilled on ice for 30 min
676 to initiate cold shock. Protein expression was induced with 0.5 mM IPTG for 18 h at
677 15°C. Cells were harvested by centrifugation, washed twice with PBS, and cell pellets
678 were flash frozen and stored at −80°C. Cell pellets were resuspended in lysis/GST
679 column buffer containing 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 mM DTT, 5% v/v
680 glycerol, and complete EDTA-free protease inhibitor cocktail (Roche). Cells were lysed
681 by a single pass at 18,000 psi through a high-pressure microfluidizer (Microfluidics
682 M110P), and the resulting lysate clarified at 30,000 × *g* for 1 h at 4°C. Clarified lysate
683 was filtered through a 0.2 µm low-protein binding syringe filter (Millex Durapore; EMD
684 Millipore) and applied to glutathione Sepharose 4b resin (Cytiva) equilibrated with GST
685 column buffer. After draining the flowthrough, the resin was washed with 50 column-
686 volumes GST column buffer. To elute the bound protein and cleave the GST tag in a
687 single step, 50 U HRV3C Protease (Novagen) was added to the column, the column

688 was sealed and incubated for 3 h at 4°C, following which, the column was drained to
689 collect the cleaved protein. The protein was diluted to 50 mM NaCl and further purified
690 using a Hi-Trap Q (Cytiva) anion exchange column equilibrated with 20 mM Tris-HCl,
691 pH 7.5, 50 mM NaCl, 1 mM DTT, and 5% v/v glycerol. The bound protein was eluted
692 using a 100–500 mM NaCl gradient in the same buffer. Peak fractions were analyzed
693 for purity by SDS-PAGE and the purest were pooled and dialyzed into storage buffer
694 containing 30 mM HEPES-KOH, pH 7.5, 100 mM potassium acetate, 3.5 mM
695 magnesium acetate, 1 mM DTT, 20% v/v glycerol. To avoid precipitation of GTSF2,
696 which has a pI = 7.3, 20 mM Tris-HCl, pH 8.8, was substituted for HEPES-KOH during
697 purification and dialysis.

698 For Figure 3f and Extended Data Fig. 5b, HEK293T cells were transfected with
699 mouse and insect GTSF1 coding sequences cloned into pcDNA5 expression vectors
700 using Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's
701 instructions. Cells were harvested 36 h later and homogenized in lysis buffer [20 mM
702 Tris-HCl (pH 7.4), 150 mM NaCl, 1.5 mM $MgCl_2$, 0.15% v/v Triton X-100, 100 μg/ml
703 RNase A (Qiagen), 0.5 mM DTT, 1× Complete EDTA-free protease inhibitor (Roche)].
704 After centrifugation at 17,000 × *g* for 20 min at 4°C, additional NaCl (final concentration:
705 0.65 M) and 1% v/v Triton X-100 (f.c.) were added to the supernatant and the lysate
706 was incubated with Streptavidin-Sepharose High Performance (Cytiva) beads at 4°C for
707 1 h. The beads were washed with wash buffer [20 mM Tris-HCl (pH 7.4), 1 M NaCl,
708 1.5 mM $MgCl_2$, 1% v/v Triton X-100, 0.5 mM DTT] five times and rinsed with lysis buffer
709 without RNase A. SBP-tagged recombinant proteins were eluted with elution buffer
710 (30 mM HEPES-KOH, pH 7.4, 100 mM potassium acetate, 2 mM magnesium acetate,
711 2.5 mM biotin, 0.5 mM DTT).

712 For Figure 3g, EmGtsf1 expression vector was transformed into BL21(DE3) cells
713 (New England Biolabs). Transformed cells were grown in LB supplemented with 1 μM
714 ZnSO4 at 37°C until OD600 ~0.6–0.8. The incubation temperature was lowered to 16°C

715 and protein expression was induced by addition of 1 mM IPTG for 4 h. Cells were
716 harvested by centrifugation and cell pellets were flash frozen in liquid nitrogen and
717 stored at −80°C for future use. Thawed cell pellets were resuspended in lysis buffer
718 (50 mM Tris, pH 8, 300 mM NaCl, 0.5 mM TCEP) and passed through a high-pressure
719 (18,000 psi) microfluidizer (Microfluidics M110P) to induce cell lysis. The resulting lysate
720 was clarified by centrifugation at 30,000 × *g* for 20 min at 4°C. Clarified lysate was
721 applied to Ni-NTA resin (Qiagen) and incubated for 1 h. The resin was extensively
722 washed with Nickel Wash Buffer (300 mM NaCl, 20 mM imidazole, 0.5 mM TCEP,
723 50 mM Tris, 8). Protein was eluted in four column volumes of Nickel Elution Buffer
724 (300 mM NaCl, 300 mM imidazole, 0.5 mM TCEP, 50 mM Tris, pH 8). TEV protease
725 was added to the eluted protein to induce cleavage and removal of the N-terminal His6
726 and MBP tags. The resulting mixture was dialyzed against Hi-Trap Dialysis Buffer
727 (300 mM NaCl, 20 mM imidazole, 0.5 mM TCEP, 50 mM Tris, pH 8) at 4°C overnight.
728 The dialyzed protein was then passed through a 5-ml Hi-Trap Chelating column (Cytiva)
729 and the unbound material collected. Unbound material was concentrated and further
730 purified by size exclusion chromatography using a Superdex 200 Increase 10/300
731 column (Cytiva) equilibrated in 50 mM Tris, pH 8, 300 mM NaCl, and 0.5 mM TCEP.
732 Peak fractions were analyzed for purity by SDS-PAGE, and the purest were pooled,
733 concentrated to 150 μM, aliquoted, and stored at −80°C.

734 **Immunodepletion and Western Blotting**

735 Approximately 300,000 FACS-purified Secondary Spermatocytes (Sp2) from wild-type
736 (C57BL/6) or 3XFLAG-HA-GTSF1-expressing (*Gtsf1*$^{KI/KI}$) mice were lysed in lysis buffer
737 (30 mM HEPES-KOH, pH 7.5, 100 mM potassium acetate, 3.5 mM magnesium acetate,
738 1 mM DTT, 0.1% v/v Triton X-100, 20% v/v glycerol, and 1× protease inhibitor cocktail
739 [1 mM 4-(2-Aminoethyl)benzenesulfonyl fluoride hydrochloride [Sigma; A8456], 0.3 μM
740 Aprotinin, 40 μM betanin hydrochloride, 10 μM. E-64 [Sigma; E3132], 10 μM leupeptin

741 hemisulfate]). Crude cytoplasmic lysate was clarified at 20,000 × *g* (S20) for 30 min at
742 4°C. The S20 protein concentration was measured using Bradford Assay. The S20 was
743 incubated with Anti-FLAG paramagnetic beads overnight at 4°C. The supernatant was
744 removed, and the beads washed thrice with lysis buffer but containing 1M potassium
745 acetate (high salt wash) and then twice with lysis buffer. Bound proteins were eluted by
746 incubation with 200 ng/ul 3XFLAG peptide for 2 h at 4°C. The extent of depletion was
747 assayed by western blotting. Briefly, the samples were resolved on SDS-PAGE and
748 transferred to nitrocellulose membrane (Amersham Protran 0.45 NC; Cytiva). The
749 membrane was probed using mouse anti-FLAG antibody (Sigma, F3165; 1:1,000 in
750 blocking buffer, Rockland Immunochemicals, MB-070) and rabbit anti-Tubulin (Cell
751 Signaling, 2144; 1:1000) for 2 hours at room temperature. After washing 5× with TBS
752 (+0.1% v/v Tween-20), the membrane was incubated with secondary antibodies -
753 donkey anti-rabbit IRDye 680RD (LI-COR, 926-68073; 1:15,000) and goat anti-mouse
754 IRDye 800CW (LI-COR, 926-32210; 1:15,000) for 30 minutes at room temperature. The
755 membrane was washed 5× with TBS (+0.1% v/v Tween-20) and imaged using Odyssey
756 Infrared Imaging System (LI-COR).

757 **Northern Blotting**

758 Northern blotting was performed as described[106]. Briefly, piRNA guide standards and
759 PIWI RISCs were first resolved on a denaturing 15% polyacrylamide gel, followed by
760 transfer to Hybond-NX (Cytiva) neutral nylon membrane by semi-dry transfer at 20 V for
761 1 h. Next, crosslinking was performed in the presence of 0.16 M EDC in 0.13 M 1-
762 methylimidazole, pH 8.0, at 60°C for 1 h. The crosslinked membrane was pre-hybridized
763 in Church's buffer (1% w/v BSA, 1 mM EDTA, 0.5 M phosphate buffer, and 7% w/v
764 SDS) at 45°C for 1 h. Radiolabeled, 5′ ^{32}P-DNA probe (25 pmol) in Church's buffer was
765 added to the membrane and allowed to hybridize overnight at 45°C, followed by five

766 washes with 1× SSC containing 0.1% w/v SDS. The membrane was air dried and
767 exposed to a storage phosphor screen.

768 **Chromatographic fractionation of the MIWI-potentiating activity**

769 Dissected animal tissues were homogenized in lysis buffer in a Dounce homogenizer
770 using 10 strokes of the loose-fitting pestle A, followed by 20 strokes of tight-fitting pestle
771 B. Lysate was clarified at 20,000 × *g*, followed by 0.2 μm filtration to yield an S20 for
772 further chromatographic purification. Lysates used without further purification were
773 directly prepared in 30 mM HEPES-KOH, pH 7.5, 100 mM potassium acetate, 3.5 mM
774 magnesium acetate, 1 mM DTT, and 20% v/v glycerol ("dialysis buffer") with 1×
775 protease inhibitor homemade cocktail; column fractions were dialyzed into this buffer
776 before assaying. Protein concentration was measured using the BCA assay.
777 Chromatography buffers were filtered prior to use.
778 For chromatography, lysate was prepared as described except using 30 mM
779 HEPES-KOH, pH 7.5, 50 mM NaCl, 1 mM DTT, 5% v/v glycerol, and protease
780 inhibitors. The lysate was applied to HiTrap SP column (Cytiva) equilibrated with the
781 lysis buffer. The column was washed, and the bound proteins eluted stepwise using
782 increasing NaCl concentrations. The NaCl content of the SP column fractions containing
783 the peak MIWI-potentiating activity was adjusted to 2 M and applied to HiTrap Phenyl
784 (Cytiva) equilibrated with column buffer containing 2 M NaCl. Bound proteins were
785 eluted stepwise using decreasing NaCl concentrations. The peak MIWI-potentiating
786 fractions elute from the HiTrap Phenyl column were pooled, concentrated (10 kDa
787 MWCO Amicon Ultra filter), and applied to Superdex 200 Increase 10-300 GL size-
788 exclusion chromatography column (bed volume ~24 ml) equilibrated with the dialysis
789 buffer but containing 5% v/v glycerol. The void volume (V_0) of the gel filtration column
790 was determined with blue dextran, and all fractions (0.5 ml each), starting from just
791 before V_0 to the end of the column (V_t) were assayed for MIWI potentiating activity. The

792 molecular weight of the potentiating activity was determined relative to molecular weight
793 markers (beta amylase, 200 kDa; alcohol dehydrogenase, 150 kDa; albumin, 66 kDa;
794 carbonic anhydrase, 29 kDa; and cytochrome C, 12.4 kDa).

795 **Zn^{2+} and Ni^{2+} immobilized-metal affinity chromatography**

796 HiTrap Chelating HP (Cytiva) columns were charged with 0.1 M $NiSO_4$ or $ZnSO_4$,
797 washed with water, and then equilibrated in column buffer (20 mM potassium phosphate
798 buffer, pH 7.5, 500 mM NaCl, 0.5 mM DTT, 5% v/v glycerol). S20 testis lysate was
799 applied to the column, the flow-through collected, and the column washed with the
800 column buffer until absorbance at 280 nm stabilized. Bound proteins were eluted in two
801 steps: first, with 20 mM potassium phosphate, pH 7.5, 2 M ammonium chloride, 0.5 mM
802 DTT, 5% v/v glycerol (elution buffer 1), and, second, with 20 mM potassium phosphate,
803 pH 7.5, 500 mM NaCl, 200 mM imidazole, pH 8.0, 0.5 mM DTT, 5% v/v glycerol (elution
804 buffer 2). The peak of each step was dialyzed into the dialysis buffer and assayed for
805 the ability to potentiate MIWI catalysis.

806 **Target cleavage assays**

807 Target RNA substrates for in vitro cleavage assays were prepared as described[74,107,108].
808 Briefly, piRNA target site-containing templates were amplified by PCR, in vitro
809 transcribed with T7 RNA polymerase, purified by urea-PAGE, and radiolabeled using α-
810 ^{32}P GTP (3000 Ci/mmol; Perkin Elmer), *S*-adenosyl methionine, and vaccinia virus RNA
811 guanylyl transferase as described[108]. Unincorporated α-^{32}P GTP was removed using a
812 G-25 spin column (Cytiva), and target RNA gel purified. In Extended Data Fig. 7a, target
813 cleavage was monitored using synthetic RNA oligonucleotides radiolabeled by ligating
814 [5′-^{32}P] cytidine 3′,5′-bisphosphate to the 3′ end of the target with T4 RNA ligase I
815 (Ambion). The [5′-^{32}P] cytidine 3′,5′-bisphosphate was prepared by incubating 1 mM
816 cytidine 3′-monophosphate (Sigma) with 312.5 pmole [γ-^{32}P] ATP (6000 Ci/mmol;

817 Perkin Elmer) with 25 U T4 polynucleotide kinase (NEB) at 37°C for 1 h, followed by
818 70°C for 30 min to inactivate the kinase.

819 Radiolabeled target (3–100 nM f.c.) was added to purified PIWI piRISC (2–8 nM),
820 plus ~1 µg tissue or sorted germ cell lysate per 10 µl reaction volume or 0.5 µM (f.c.) of
821 purified GTSF protein. At the indicated times, an aliquot of a master reaction was
822 quenched in 4 volumes 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 25 mM EDTA, 1% w/v
823 SDS, then proteinase K (1 mg/ml f.c.) was added and incubated at 45°C for 15 min. An
824 equal volume of urea loading buffer (8 M urea, 25 mM EDTA) was added to the reaction
825 time points, heated at 95°C for 2 min, and resolved by 7–10% denaturing PAGE. Gels
826 were dried, exposed to a storage phosphor screen, and imaged on a Typhoon FLA
827 7000 (GE).

828 The raw image file was used to quantify the substrate and product bands,
829 corrected for background. Data were fit to the reaction scheme

830 $$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \overset{k_2}{\rightarrow} EP \overset{k_3}{\rightarrow} E + P$$

831 using the burst-and-steady-state equation, Equation 1.

832 $$[P] = f(t) = [E]([k_2/(k_2 + k_3)]^2 \times (1 - e^{-[k2+k]t}) + [k_2k_3/(k_2 + k_3)]t) \qquad \text{(Equation 1)}$$

833 The time-dependence of product formation corresponded to a pre-steady-state
834 exponential burst ($k_{burst} = k_2 + k_3$) followed by a linear steady-state phase, described by
835 k_{cat}, *where* $k_{cat} = k_{ss} = k_2k_3/(k_2 + k_3)$.

836 The affinity (K_D) of wild-type or GTSF1^{W98A,W107A,W112A} for MIWI piRISC-target
837 ternary complex and the maximum observable rate (k_{pot}) were estimated by measuring
838 the pre-steady-state rate of target cleavage (k_{burst}) of MIWI at increasing concentrations
839 of GTSF1 (1–5,000 nM) and fitting the data to Equation 2 (ref. 109:

840 $$k_{burst} = k_{pot}\,[\text{GTSF1}]/(K_D + [\text{GTSF1}]) \qquad \text{(Equation 2).}$$

841 **FLAG-Siwi, MIWI, and MILI target-cleavage assays**

842 FLAG-tagged BmSiwi, MIWI, or MILI-expressing vectors were transfected into BmN4
843 cells with X-tremeGENE HP DNA Transfection Reagent (Sigma). Preparation of loading
844 lysates and single-stranded RNA loading were as described[110]. After loading with
845 synthetic guide RNAs in the cell lysate, piRISC was immunoprecipitated with anti-FLAG
846 antibody (Sigma) conjugated to Dynabeads protein G superparamagnetic beads
847 (Thermo Fisher Scientific). The beads were washed five times with lysis buffer
848 containing 30 mM HEPES-KOH, pH 7.4, 100 mM potassium acetate, 2 mM magnesium
849 acetate, 0.5 mM DTT, and 0.1% v/v Empigen. Target cleavage assays were performed
850 at 25°C for 3.5 h in 8 μl reaction containing 2.4 μl "40×" reaction mix[108]. 100 nM
851 recombinant GTSF1 protein and 0.5 nM ^{32}P cap-radiolabeled 28 nt target RNA as
852 described[24].

853 **BmAgo3 target-cleavage assays**

854 Naive BmN4 cells were resuspended in lysis buffer (30 mM HEPES-KOH, pH 7.4,
855 100 mM potassium acetate, 2 mM magnesium acetate, 0.05% v/v Triton X-100, 0.5 mM
856 DTT, 1× Complete EDTA-free protease inhibitor [Roche]) and lysed using a Dounce
857 homogenizer. After centrifugation at 17,000 × *g* for at 4°C 20 min, the supernatant was
858 incubated at 4°C for 1 h with anti-BmAgo3 antibody[23] conjugated to Dynabeads
859 protein G (Thermo Fisher Scientific). The superparamagnetic beads were washed five
860 times with wash buffer (30 mM HEPES-KOH, pH 7.4, 100 mM potassium acetate, 2 mM
861 magnesium acetate, 0.5% v/v Empigen, 0.5% v/v Triton X-100, 0.5 mM DTT) and rinsed
862 with lysis buffer. Target cleavage assays were performed at 25°C for 2.5 h in 8 μl
863 reactions containing 2.4 μl "40×" reaction mix, 55 nM recombinant protein, and 0.5 nM
864 ^{32}P cap-radiolabeled target 1 (complementary to endogenous piRNA #1) or target 2
865 (complementary to endogenous piRNA #2), prepared by in vitro transcription as
866 described[111].

867 **EfPiwi target cleavage assays**

868 EfPiwi piRISC was purified as described[64]. Purified EfPiwi piRISC (100 nM, final
869 concentration) was incubated at 37°C with a 5′-^{32}P-labeled target RNA complementary
870 to g2-g21 (10 nM, final concentration) in reaction buffer composed of 20 mM Tris pH
871 8.0, 100 mM NaCl, 2 mM $MgCl_2$, 2 mM $MnCl_2$, 0.5 mM TCEP, and 0.1 mg/mL baker's
872 yeast tRNA. Where indicated, purified *Ephydatia muelleri* GTSF1 (500 nM, final
873 concentration) was also included. Target cleavage was stopped at indicated times by
874 mixing aliquots of each reaction with an equal volume of denaturing gel loading buffer
875 (98% w/v formamide, 0.025% xylene cyanol, 0.025% w/v bromophenol blue, 10 mM
876 EDTA pH 8.0). Intact and cleaved target RNAs were resolved by denaturing PAGE
877 (15%) and visualized by phosphorimaging. Quantification of signal was performed using
878 ImageQuant TL (GE Healthcare).

879 **Mouse germ cell purification**

880 Germ cells from mouse testes were sorted and purified as described[16,47]. Briefly, freshly
881 dissected mouse testes were decapsulated with 0.4 mg/ml collagenase type IV
882 (Worthington) in 1× Gey's balanced salt solution (GBSS) at 33°C for 15 min. The
883 separated seminiferous tubules were treated with 0.5 mg/ml trypsin and 1 μg/ml
884 DNase I in 1× GBSS at 33°C for 15 min. Trypsin was then inactivated by adding 7.5%
885 v/v fetal bovine serum (FBS). The cell suspension was filtered through a 70-μm cell
886 strainer, and cells pelleted at 300 × *g* at 4°C for 10 min. Cell staining was performed at
887 33°C for 15 min with 5 μg/ml Hoechst 33342 prepared in 1× GBSS, 5% v/v FBS, and
888 1 μg/ml DNase I, then the cells were treated with 0.2 μg/ml propidium iodide, followed
889 by final pass through a 40-μm cell strainer before sorting. Sorted cells were pelleted at
890 100 × *g* for 5 min, the buffer removed, and the cell pellets flash frozen and stored at
891 −80°C. Cell lysates were prepared as described for HEK293T cells. Protein

892 concentration was estimated using the BCA assay, and an equal amount of total protein
893 from each cell type was used to assay for the ability to potentiate MIWI catalysis.

894 **Analysis of RNA-seq data**

895 Publicly available datasets[45,112–114] were analyzed. rRNA reads were removed using
896 Bowtie 2.2.5 with default parameters[115]. After rRNA removal, the remaining reads were
897 mapped to corresponding genomes (mouse, mm10; rat, rn6; macaque, rheMac8;
898 human, hg19) using STAR. 2.3 with default parameters that allowed ≤ 2 mismatches
899 and 100 mapping locations (ref. 116). Mapped results were generated in SAM format,
900 duplicates removed and translated to BAM format using SAMtools 1.8 (ref. 117).
901 HTSeq 0.9.1 with default parameters was used to count uniquely mapping reads[118]
902 (steady-state transcript abundance was reported in reads per kilobase per million
903 uniquely mapped reads [RPKM])

904 **Methylarginine analysis**

905 Recombinant MIWI was immunopurified and resolved by electrophoresis on a 4–20%
906 gradient SDS-polyacrylamide gel. The gel was fixed, stained with Coomassie G-250
907 (Simply Blue, Invitrogen), and the recombinant MIWI band excised and analyzed at the
908 UMASS Mass Spectrometry Core. Gel slices were chopped into ~1 mm^2 pieces, 1 ml
909 water added, followed by 20 µl 45 mM DTT in 250 mM ammonium bicarbonate.
910 Samples were incubated at 50°C for 30 min, cooled to room temperature, then 20 µl
911 100 mM iodoacetamide (IAA) was added and incubated for 30 min. The solution was
912 removed, and the gel pieces were three times washed with 1 ml water, 1 ml 50 mM
913 ammonium bicarbonate:acetonitrile (1:1), quenched with 200 µl acetonitrile, and dried in
914 a SpeedVac. Gel pieces were then rehydrated in 50 ul 50 mM ammonium bicarbonate
915 containing 4 ng/µl trypsin (Promega, Madison, WI) and 0.01% proteaseMAX (Promega)
916 and incubated at 37°C for 18 h. Supernatants were collected and extracted with 200 µl

917 an 80:20 solution of acetonitrile: 1% (v/v) formic acid in water. Supernatants were then
918 combined, and the peptides lyophilized in a SpeedVac and re-suspended in 25 µl 5%
919 acetonitrile, 0.1% (v/v) formic acid and subject to mass spectrometry analysis. Data
920 were acquired using a NanoAcquity UPLC (Waters Corporation, Milford, MA) coupled to
921 an Orbitrap Fusion Lumos Tribrid (Thermo Fisher Scientific, Waltham, MA) mass
922 spectrometer. Peptides were trapped and separated using an in-house 100 µm I.D.
923 fused-silica pre-column (Kasil frit) packed with 2 cm ProntoSil (Bischoff
924 Chromatography, DE) C18 AQ (200 Å, 5 µm) media and configured to an in-house
925 packed 75 µm I.D. fused-silica analytical column (gravity-pulled tip) packed with 25 cm
926 ProntoSil (Bischoff; 100 Å, 3 µm) media. Mobile phase A was 0.1 % (v/v) formic acid in
927 water; mobile phase B was 0.1 % (v/v) formic acid in acetonitrile. Following a 3.8 µl
928 sample injection, peptides were trapped at flow rate of 4 µl/min with 5% B for 4 min,
929 followed by gradient elution at a flow rate of 300 nl/min from 5–35% B in 90 min (total
930 run time, 120 min). Electrospray voltage was delivered by liquid junction electrode
931 (1.5 kV) located between the columns and the transfer capillary to the mass
932 spectrometer was maintained at 275°C. Mass spectra were acquired over *m/z* 300–
933 1750 Da with a resolution of 120,000 (*m/z* 200), maximum injection time of 50 ms, and
934 an AGC target of 400,000. Tandem mass spectra were acquired using data-dependent
935 acquisition (3 s cycle) with an isolation width of 1.6 Da, HCD collision energy of 30%,
936 resolution of 15,000 (*m/z* 200), maximum injection time of 22 ms, and an AGC target of
937 50,000.

938 Raw data were processed using Proteome Discoverer 2.1.1.21 (Thermo Fisher
939 Scientific), and the database search performed by Mascot 2.6.2 (Matrix Science,
940 London, UK) using the Swiss-Prot human database (download 04/09/2019). Search
941 parameters were: semi-tryptic digestion with up to two missed cleavages; precursor
942 mass tolerance 10 ppm; fragment mass tolerance 0.05 Da; peptide N-terminal
943 acetylation, cysteine carbamidomethylation, methionine oxidation, N-terminal glutamine

944 to pyroglutamate conversion, arginine methylation and arginine demethylation were
945 specified as variable modifications. Peptide and protein validation and annotation was
946 done in Scaffold 4.8.9 (Proteome Software, Portland, OR) employing Peptide Prophet[119]
947 and Protein Prophet[120]. Peptides were filtered at 1% FDR, while protein identification
948 threshold was set to greater than 99% probability and with a minimum of two identified
949 peptides per protein. Only arginine modification sites detected in all three replicates
950 from separate immunoprecipitation experiments are reported in the figure.

951 **Statistics and reproducibility**

952 Figs. 1b, 3, 4 and Extended Data Figs. 4a, 5c show mean ± SD for three independent
953 trials. Figs. 2a, b, Extended Data Figs. 1a–c and 1h (three independent trials) and
954 Figs. 2e, f, and Extended Data Figs. 1f–g, 2b, 3a, 3c (two independent trials) and
955 Extended Data Figs. 4b, 4c show representative data. The experiments in Figs. 1c, 5a,
956 and Extended Data Figs. 1d–e, 2, 3b, 3d–h, 5a, 7a–b were performed once. Protein
957 sequences were aligned using Clustal Omega; unrooted tree was constructed using
958 Randomized Axelerated Maximum Likelihood with default parameters[121] and visualized
959 in Interactive Tree Of Life[122].

960 **Reporting summary**

961 Further information on research design is available in the Nature Research Reporting
962 Summary linked to this paper.

963 **Data availability**

964 All data are available from the authors upon request.

965 102. Kawaoka, S. et al. The *Bombyx* ovary-derived cell line endogenously expresses
966 PIWI/PIWI-interacting RNA complexes. *RNA* **15**, 1258-1264 (2009).

967 103. Aoshima, K., Baba, A., Makino, Y. & Okada, Y. Establishment of alternative culture
968 method for spermatogonial stem cells using knockout serum replacement. *PLoS*
969 *One* **8**, e77715 (2013).

970 104. Kenny, N. J. et al. Tracing animal genomic evolution with the chromosomal-level
971 assembly of the freshwater sponge Ephydatia muelleri. *Nat Commun* **11**, 3676
972 (2020).

973 105. Flores-Jasso, C. F., Salomon, W. E. & Zamore, P. D. Rapid and specific
974 purification of Argonaute-small RNA complexes from crude cell lysates. *RNA* **19**,
975 271-279 (2013).

976 106. Pall, G. S. & Hamilton, A. J. Improved northern blot method for enhanced detection
977 of small RNA. *Nat Protoc* **3**, 1077-1084 (2008).

978 107. Wee, L., Flores-Jasso, C., Salomon, W. & Zamore, P. Argonaute Divides Its RNA
979 Guide Into Domains with Distinct Functions and RNA-Binding Properties. *Cell* **151**,
980 1055-1067 (2012).

981 108. Haley, B., Tang, G. & Zamore, P. D. In vitro analysis of RNA interference in
982 *Drosophila melanogaster*. *Methods* **30**, 330-336 (2003).

983 109. Carlson, K. D., Johnson, R. E., Prakash, L., Prakash, S. & Washington, M. T.
984 Human DNA polymerase kappa forms nonproductive complexes with matched
985 primer termini but not with mismatched primer termini. *Proc. Natl. Acad. Sci. U S A*
986 **103**, 15776-15781 (2006).

987 110. Kawaoka, S. et al. Zygotic amplification of secondary piRNAs during silkworm
988 embryogenesis. *RNA* **17**, 1401-1407 (2011).

989 111. Yoda, M. et al. ATP-dependent human RISC assembly pathways. *Nat. Struct. Mol.*
990 *Biol.* **17**, 17-23 (2010).

991 112. Merkin, J., Russell, C., Chen, P. & Burge, C. B. Evolutionary dynamics of gene and
992 isoform regulation in Mammalian tissues. *Science* **338**, 1593-1599 (2012).

993 113. Yu, T. et al. Long first exons and epigenetic marks distinguish conserved
994 pachytene piRNA clusters from other mammalian genes. *Nat Commun* **12**, 73
995 (2021).

996 114. Gainetdinov, I. et al. Terminal Modification, Sequence, and Length Determine
997 Small RNA Stability in Animals. *bioRxiv* 2020.09.08.287979 (2020).

998 115. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat*
999 *Methods* **9**, 357-359 (2012).

1000 116. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-
1001 21 (2013).

1002 117. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**,
1003 2078-2079 (2009).

1004 118. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-
1005 throughput sequencing data. *Bioinformatics* **31**, 166-169 (2015).

1006 119. Keller, A., Nesvizhskii, A. I., Kolker, E. & Aebersold, R. Empirical statistical model
1007 to estimate the accuracy of peptide identifications made by MS/MS and database
1008 search. *Anal. Chem.* **74**, 5383-5392 (2002).

1009 120. Nesvizhskii, A. I., Keller, A., Kolker, E. & Aebersold, R. A statistical model for
1010 identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**, 4646-4658
1011 (2003).

1012 121. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis
1013 of large phylogenies. *Bioinformatics* **30**, 1312-1313 (2014).

1014 122. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display
1015 and annotation of phylogenetic and other trees. *Nucl. Acids Res.* **44**, W242-W245
1016 (2016).

1017 **Acknowledgements** We are grateful to Ken-Edwin Aryee (Greiner/Bremh Lab) for
1018 providing pScalps_Puro, psPAX2, pMD2.G plasmids and for guidance on lentiviral
1019 transduction; Pei-Hsuan Wu for help with mouse testis germ cell FACS and for
1020 providing sorted cells; Katharine Cecchini for mouse tissues; Elaine Norowski
1021 (Greiner/Bremh labs) for rat testes; Kyle Orwig (University of Pittsburgh) for rhesus
1022 macaque testes; Paul Albosta and Cindy Tipping for guidance with *T. ni* and *D.*
1023 *melanogaster* dissection; Leemor Joshua-Tor (Cold Spring Harbor Laboratory) for
1024 sharing data prior to publication; Tiffanie Gardner for help formatting the manuscript; the
1025 UMASS Flow Cytometry and Mass Spectrometry cores; and members of the Zamore
1026 lab for discussions and comments on the manuscript. This work was supported in part
1027 by NIGMS grants R37 GM062862 and R35 GM136275 (P.D.Z.) and R35 GM127090
1028 (I.J.M) and JSPS KAKENHI grants 18H05271 (Y.T.) and 19K06484 (N.I.).

1029 **Author contributions** A.A., P.D.Z., I.J.M., and Y.T. conceived and designed the
1030 experiments. A.A., S.M.B., N.I., T.A.A., D.M.O., C.A., performed the experiments.
1031 D.M.O. and I.G. provided and analyzed the sequencing data. P.D.Z. supervised the
1032 research. A.A. and P.D.Z. wrote the manuscript. All the authors discussed the results
1033 and approved the manuscript.

1034 **Competing interests** The authors declare no competing interests.

1035 **Additional information**

1036 **Supplementary information** is available for this paper at .

1037 **Correspondence and requests for materials** should be addressed to P.D.Z.

1038

1039 **Figure legends**

1040 **Figure 1 | A component of mouse testis lysate potentiates piRNA-directed target**
1041 **RNA cleavage by MIWI.**

1042 **a,** Strategy for programming recombinant MIWI with synthetic piRNA. **b,** Top**:**
1043 representative denaturing polyacrylamide gel image for target RNA cleavage by MIWI
1044 piRISC with and without added testis lysate. Bottom: product formed as a function of
1045 time by MIWI piRISC (mean ± SD, *n* =3). Rate constants were determined by fitting the
1046 data to the burst-and-steady-state equation (Equation 1 in Methods). $[E]_{active}$: the
1047 apparent concentration of active MIWI piRISC estimated from data fitting; k_{burst}: the pre-
1048 steady-state rate; k_{ss}: the steady-state rate. **c,** Target RNA cleavage by MIWI piRISC in
1049 the presence of lysate from either whole testis or FACS-purified germ cells. Spg,
1050 spermatogonia; PS, pachytene spermatocytes; DS, diplotene spermatocytes; SpII,
1051 secondary spermatocytes; Sptd, spermatids.

1052 **Figure 2 | Testis protein GTSF1 potentiates target RNA cleavage by MIWI piRISC.**

1053 **a,** Scheme to purify the MIWI-potentiating factor from mouse testis lysate. **b,** Target
1054 cleavage assay to estimate the apparent molecular weight of the MIWI-potentiating
1055 activity by size-exclusion chromatography (SEC). Arrowheads indicate the peak
1056 concentration of the molecular weight standards, and their peak elution volumes. V_0:
1057 void volume. **c,** The properties of the MIWI-potentiating activity compared to those of
1058 GTSF1. **d,** Strategy for creating knock-in mouse expressing 3XFLAG-HA-GTSF1
1059 (*Gtsf1*$^{KI/KI}$). Arrows represent sequencing primers used to validate the insertion. The
1060 exon encodes an incomplete Isoleucine codon, which is completed upon splicing. **e,**
1061 Western Blotting to show immuno-depletion of 3XFLAG-HA-GTSF1 from secondary
1062 spermatocyte lysate using anti-FLAG paramagnetic beads. I, Input; S, Supernatant; E,
1063 3XFLAG peptide eluate. **f,** Product formed by MIWI piRISC in the presence of the

1064 indicated components. Abbreviations as in **e**. Numbers below the gel indicate relative
1065 product formed in each condition.

1066 **Figure 3 | GTSF1 paralogs can distinguish between MIWI and MILI.**

1067 **a,** Representative denaturing polyacrylamide gel image showing that purified GTSF1
1068 recapitulates the effect of testis lysate on MIWI catalysis. **b,** Product generated as a
1069 function of time (mean ± SD, n = 3). Data were fit to the burst-and-steady-state
1070 equation. Data for MIWI RISC alone are from Fig. 1. **c,** Representative denaturing
1071 polyacrylamide gel images of the assay to test GTSF1 mutants and paralogs in target
1072 cleavage by MIWI or MILI. **d,** Product formed as a function of time (mean ± SD, n = 3).
1073 **e,** Observed pre-steady-state first order rate constants of target cleavage by MIWI
1074 piRISC in the presence of either wild-type or PIWI-interacting mutant
1075 GTSF1^{W98A,W107A,W112A} were plotted as a function of GTSF1 concentration. Dissociation
1076 constant (K_D) and maximum potentiation rate (k_{pot}) were determined by fitting each
1077 replicate individually to Equation 2 (see Methods). The curve corresponds to the mean
1078 K_D and k_{pot} of the three replicates (n = 3). **f,** Representative target cleavage assay using
1079 the indicated PIWI protein in the presence of different GTSF1 orthologs (100 nM).
1080 Numbers below the gel lanes report fraction target cleaved (mean ± SD; n = 3). **g,**
1081 Single-turnover rate of target cleavage by EfPiwi piRISC in the presence or absence of
1082 EmGtsf1 (mean ± SD, n = 3).

1083 **Figure 4 | Cleavage by MIWI piRISC is sensitive to the extent of complementarity**
1084 **between the piRNA and target.**

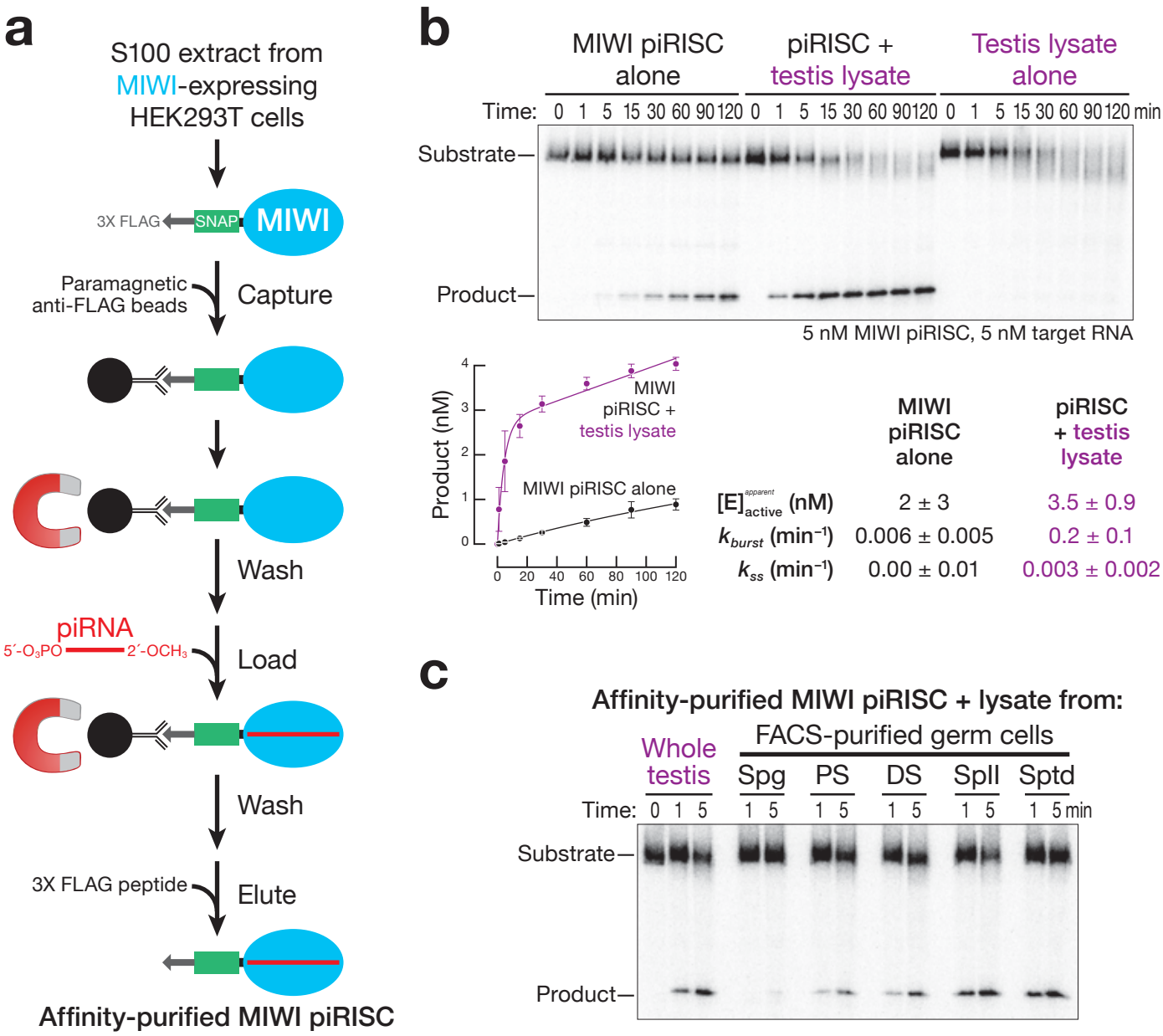
1085 **a,** MIWI piRISC target cleavage in the presence of GTSF1 for targets with increasing
1086 complementarity to synthetic piRNA guide. Top: multiple-turnover conditions; bottom:
1087 single-turnover conditions. All reactions contained saturating amounts of GTSF1. **b,**
1088 Target cleavage assay using targets complementary to piRNA guide nucleotides g2–

1089 g16 or g2–g30 and MIWI or MILI loaded with piRNAs of the indicated lengths, with or
1090 without GTSF1 (mean ± SD, *n* = 3). **c,** Absolute (mean ± SD, *n* = 3) and relative pre-
1091 steady-state and steady-state rates of cleavage of the g2–g30 target by MIWI loaded
1092 with piRNAs of the indicated lengths in the presence of GTSF1.

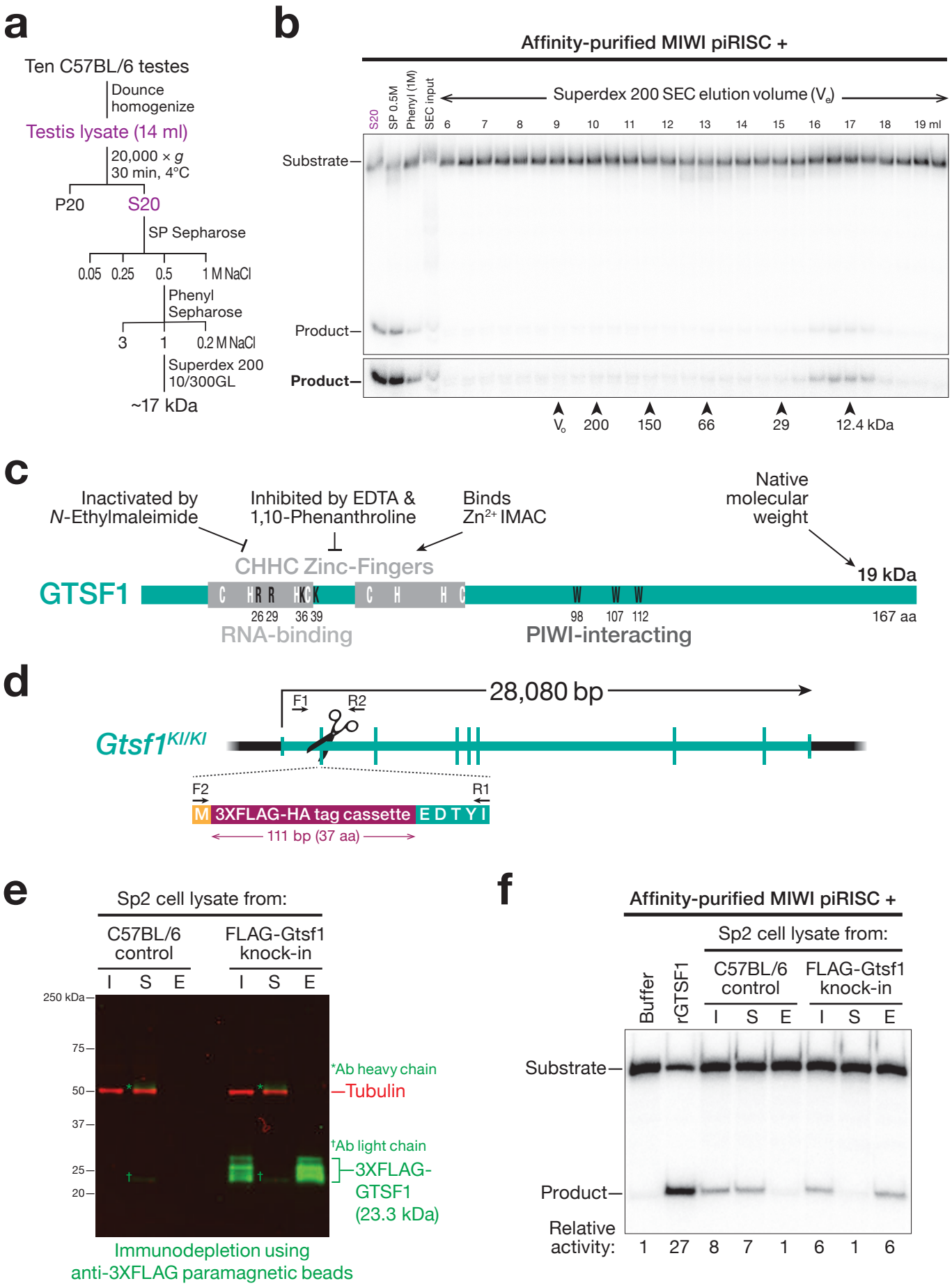
1093 **Figure 5 | A model for the function of guide length and GTSF1 in target cleavage**
1094 **by PIWI proteins.**

1095 **a,** The energy of base pairing, estimated by standard nearest-neighbor methods vs. pre-
1096 and steady-state rates of GTSF1-potentiated target cleavage (mean ± SD, *n* = 3),
1097 directed by piRNAs of different lengths loaded into MIWI. The same target, which was
1098 fully complementary to each piRNA, was used in all experiments. **b,** piRNA-directed,
1099 MIWI-catalyzed target cleavage is envisioned to require two sequential conformational
1100 changes: (1) a target-dependent conformational change in piRISC (E_{PAZ}) in which the
1101 piRNA 3′ end leaves the PAZ domain, allowing extensive base pairing of the piRNA with
1102 the target substrate (S); and (2) GTSF1-dependent conversion of this piRISC pre-
1103 catalytic state (E_C) to the fully competent catalytic state ($E_{C'}$). P: cleaved target products.
1104 **c,** Proposed effects of different piRNA lengths, extents of guide:target complementarity,
1105 and GTSF1 on the forward and reverse rates of the two conformational rearrangements.
1106 The wide, central cleft like that observed in the cryo-EM structure of *Ephydatia fluviatilis*
1107 Piwi-A[64] is envisioned to allow the central region of a 30 nt piRNA to be mobile and
1108 exposed to solvent when its 3′ end is secured to the PAZ domain (upper right).

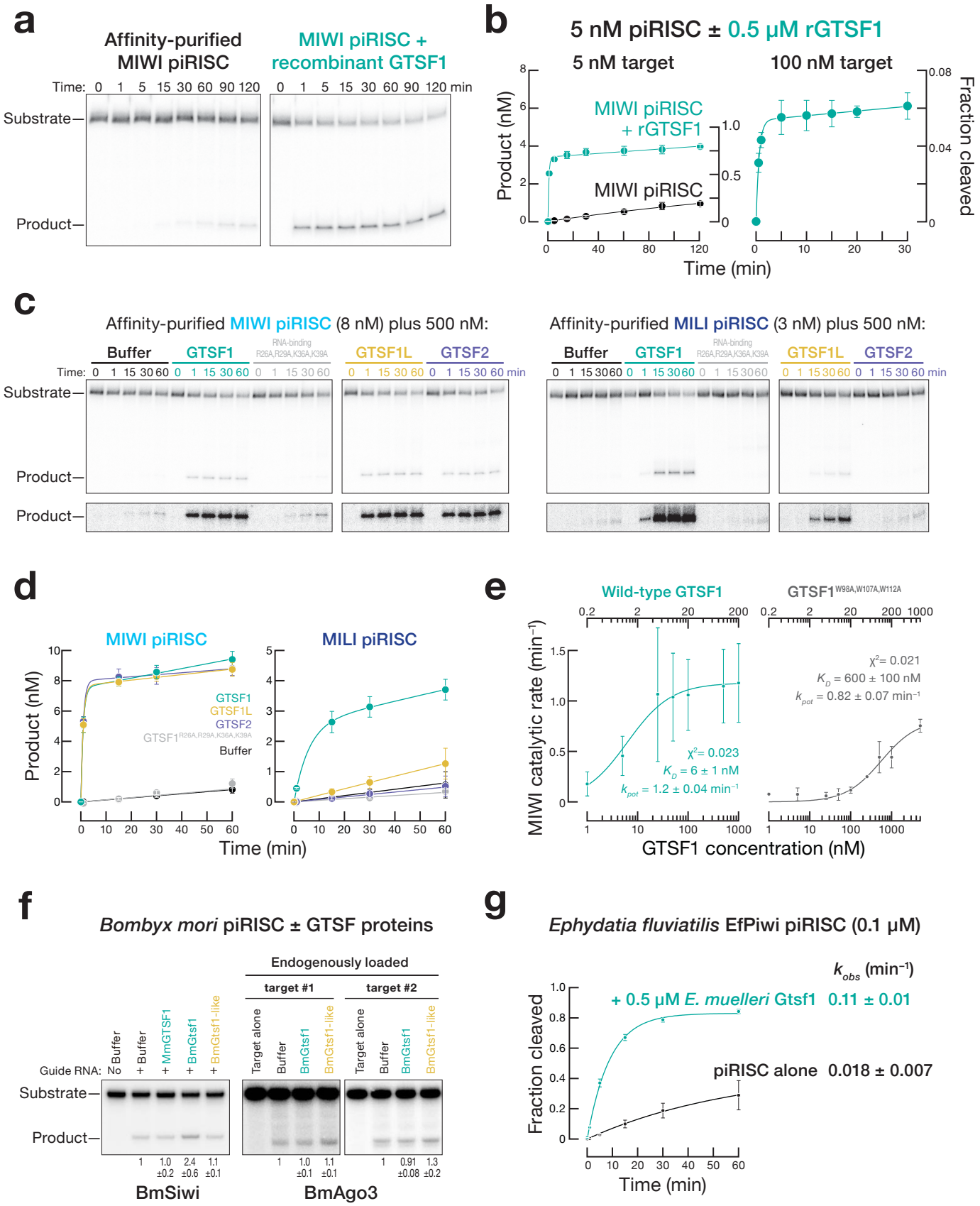
Arif et al., Figure 1



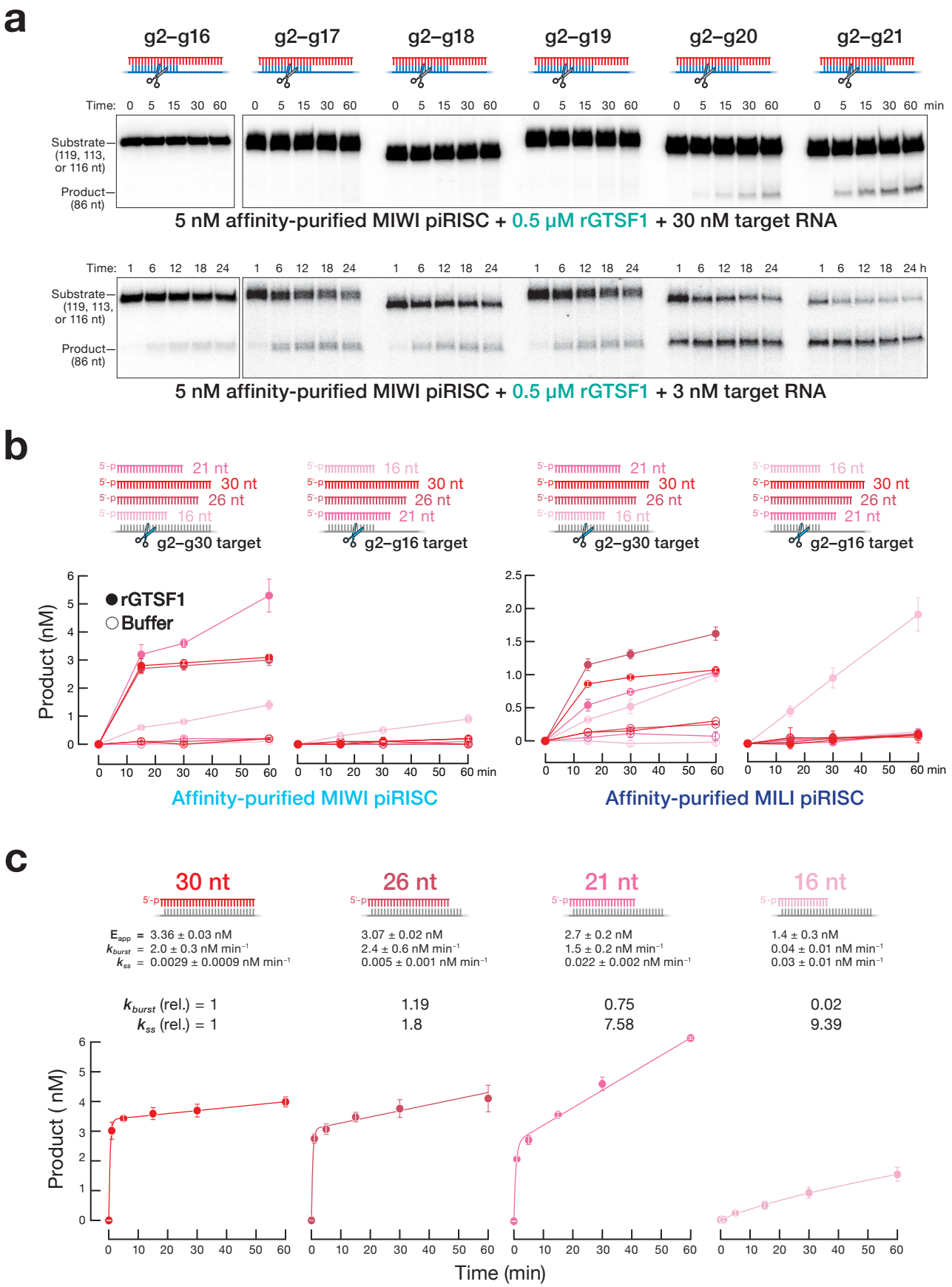
Arif et al., Figure 2



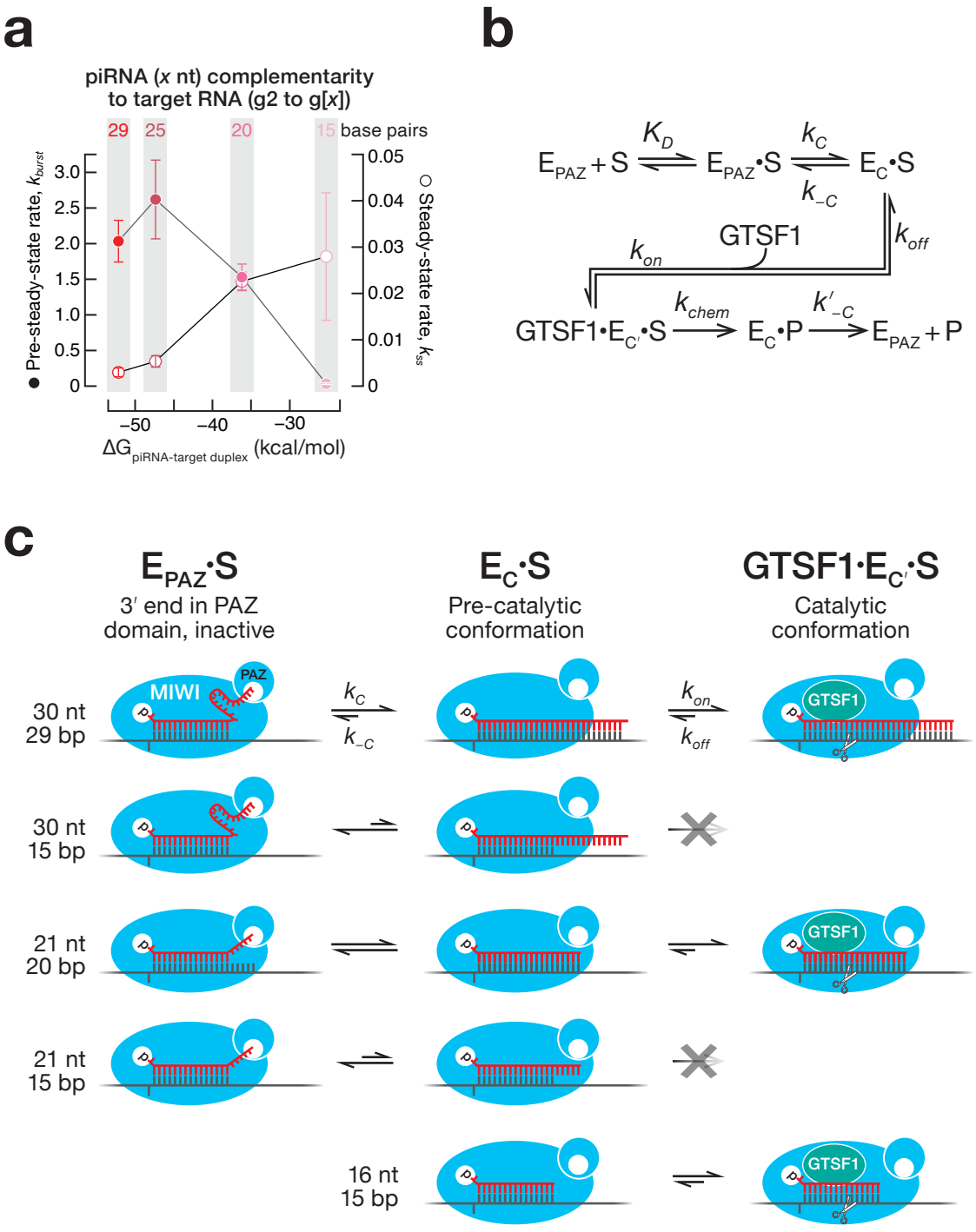
Arif et al., Figure 3



Arif et al., Figure 4



Arif et al., Figure 5



Extended Data Table 1. Pairwise comparison of percent identity or similarity for domain sequence of GTSF1-related proteins.

			Mus musculus (Mm)			*Rattus norvegicus (Rn)*			*Macaca mulatta (Mmul)*		*Trichoplusia ni (Tn)*		*Bombxy mori (Bm)*		*Drosophila melanogaster (Dm)*
			GTSF1	GTSF1L	GTSF2	GTSF1	GTSF1L	GTSF2	GTSF1	GTSF1L	TnGtsf1	TnGtsf1l	BmGtsf1	BmGtsf1l	Asterix
Zinc finger 1	***Mm***	GTSF1	100	50	54.2	100	50	54.2	100	54.2	45.8	54.2	45.8	33.3	37.5
		GTSF1L	62.5	100	73.9	50	100	73.9	50	95.7	41.7	52.2	37.5	56.5	39.1
		GTSF2	66.7	87	100	54.2	73.9	100	54.2	78.3	33.3	43.5	33.3	56.5	34.8
	Rn	GTSF1	100	62.5	66.7	100	50	54.2	100	54.2	45.8	54.2	45.8	33.3	37.5
		GTSF1L	62.5	100	87	62.5	100	73.9	50	95.7	41.7	52.2	37.5	56.5	39.1
		GTSF2	66.7	87	100	66.7	87	100	54.2	78.3	33.3	43.5	33.3	56.5	34.8
	Mmul	GTSF1	100	62.5	66.7	100	62.5	66.7	100	54.2	45.8	54.2	45.8	33.3	37.5
		GTSF1L	66.7	95.7	91.3	66.7	95.7	91.3	66.7	100	37.5	47.8	37.5	52.2	34.8
	Tn	Gtsf1	54.2	54.2	54.2	54.2	54.2	54.2	54.2	50	100	41.7	83.3	45.8	37.5
		Gtsf1-like	66.7	65.2	52.2	66.7	65.2	52.2	66.7	60.9	62.5	100	41.7	43.5	39.1
	Bm	Gtsf1	58.3	45.8	54.2	58.3	45.8	54.2	58.3	50	87.5	54.2	100	37.5	37.5
		Gtsf1-like	54.2	65.2	65.2	54.2	65.2	65.2	54.2	60.9	66.7	73.9	58.3	100	52.2
	Dm	Asterix	58.3	56.5	56.5	58.3	56.5	56.5	58.3	52.2	66.7	69.6	62.5	73.9	100
Zinc finger 2	***Mm***	GTSF1	100	39.1	39.1	100	39.1	39.1	100	39.1	43.5	43.5	39.1	30.4	39.1
		GTSF1L	65.2	100	78.3	39.1	100	78.3	39.1	82.6	39.1	30.4	34.8	30.4	26.1
		GTSF2	56.5	91.3	100	39.1	78.3	100	39.1	82.6	34.8	39.1	34.8	39.1	26.1
	Rn	GTSF1	100	65.2	56.5	100	39.1	39.1	100	39.1	43.5	43.5	39.1	30.4	39.1
		GTSF1L	65.2	100	91.3	65.2	100	78.3	39.1	82.6	39.1	30.4	34.8	30.4	26.1
		GTSF2	56.5	91.3	100	56.5	91.3	100	39.1	82.6	34.8	39.1	34.8	39.1	26.1
	Mmul	GTSF1	100	65.2	56.5	100	65.2	56.5	100	39.1	43.5	43.5	39.1	30.4	39.1
		GTSF1L	52.2	87	87	52.2	87	87	52.2	100	34.8	39.1	39.1	30.4	26.1
	Tn	Gtsf1	56.5	52.2	52.2	56.5	52.2	52.2	56.5	52.2	100	43.5	82.6	34.8	26.1
		Gtsf1-like	52.2	47.8	56.5	52.2	47.8	56.5	52.2	52.2	52.2	100	43.5	60.9	30.4
	Bm	Gtsf1	52.2	47.8	52.2	52.2	47.8	52.2	52.2	56.5	91.3	52.2	100	39.1	26.1
		Gtsf1-like	60.9	47.8	47.8	60.9	47.8	47.8	60.9	39.1	65.2	82.6	60.9	100	26.1
	Dm	Asterix	56.5	60.9	56.5	56.5	60.9	56.5	56.5	60.9	60.9	65.2	56.5	47.8	100
Carboxy terminal domain	***Mm***	GTSF1	100	7.3	8.3	100	7.3	10.4	88.5	7.3	7.3	6.2	8.3	6.2	8
		GTSF1L	20.8	100	32.6	7.3	93.2	31.5	9.4	60	4.3	4.4	5.4	4.4	5.5
		GTSF2	28.1	54.3	100	8.3	32.6	89	9.4	34.8	2.2	4.4	4.3	5.5	10.1
	Rn	GTSF1	100	20.8	28.1	100	7.3	10.4	88.5	7.3	7.3	6.2	8.3	6.2	8
		GTSF1L	20.8	97.7	54.3	20.8	100	31.5	9.4	58.9	3.2	4.4	4.3	4.4	5.5
		GTSF2	28.1	53.3	93.4	28.1	53.3	100	11.5	33.7	2.2	4.4	4.3	5.5	9.2
	Mmul	GTSF1	92.7	19.8	28.1	92.7	19.8	28.1	100	8.3	8.3	8.3	9.4	7.3	7.1
		GTSF1L	24	66.7	53.3	24	68.9	51.1	24	100	6.5	1.1	9.7	2.2	4.6
	Tn	Gtsf1	20.8	20.4	17.4	20.8	20.4	16.3	16.7	19.4	100	6.2	60	5	8.3
		Gtsf1-like	11.5	8.9	8.8	11.5	7.8	8.8	12.5	6.7	7.5	100	5	44.4	4.6
	Bm	Gtsf1	27.1	17.2	21.7	27.1	19.4	20.7	22.9	22.6	71.2	11.2	100	7.5	8.3
		Gtsf1-like	11.5	12.2	14.3	11.5	11.1	13.2	12.5	7.8	6.2	60	10	100	1.8
	Dm	Asterix	17.9	15.6	18.3	17.9	14.7	17.4	18.8	14.7	18.3	10.1	17.4	11	100

Extended Data Table 2.

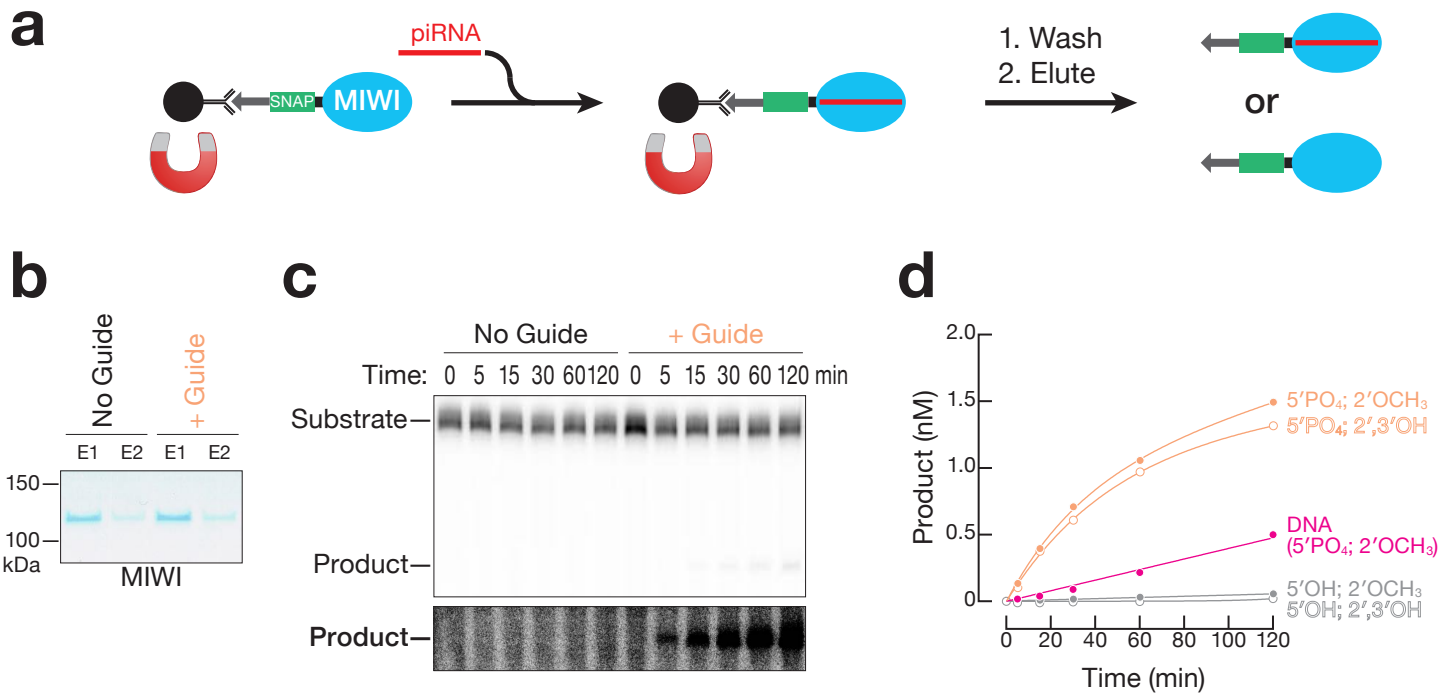
		GTSF1	GTSF1 PIWI-interaction mutant	GTSF1 RNA-binding mutant	GTSF1L	GTSF2	Buffer
MIWI	k_{burst} (min^{-1})	1.1 ± 0.1	0.67 ± 0.08	< LQ	1.1 ± 0.1	1.1 ± 0.1	< LQ
	k_{rel}	1	0.6	—	1	1	—
MILI	k_{burst} (min^{-1})	0.17 ± 0.03	0.04 ± 0.01	< LQ	0.02 ± 0.02	< LQ	< LQ
	k_{rel}	1	0.2	—	0.06	—	—

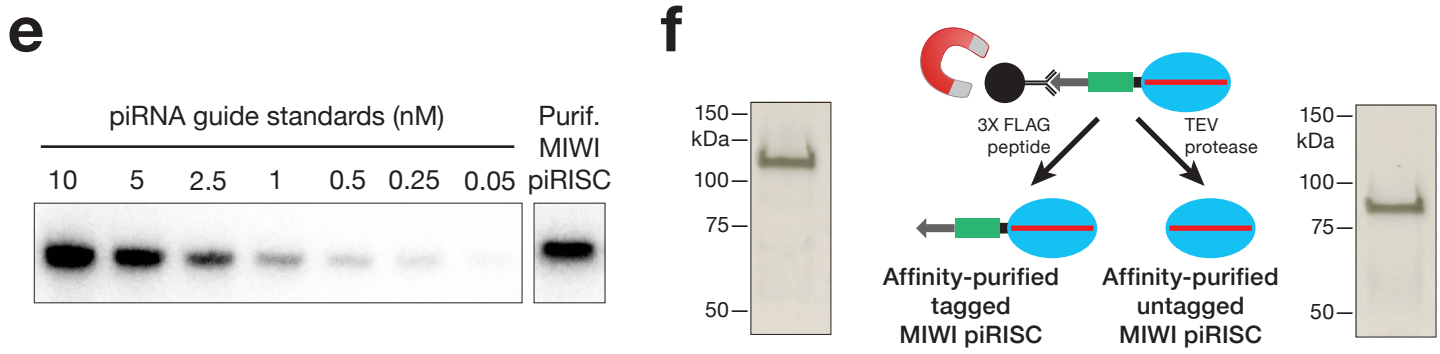
Extended Data Figures and Legends

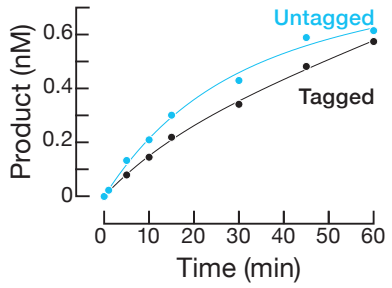
Extended Data Figure 1 | Recombinant MIWI programmed with guide piRNA cleaves complementary target RNA.

a, Purification of apo-MIWI and of MIWI loaded with a synthetic piRNA guide (piRISC). **b,** Coomassie-stained SDS-PAGE of purified apo-MIWI and MIWI piRISC. E1, E2: first and second eluates from sequential incubation of the paramagnetic beads with 3XFLAG peptide. E1 was used for the cleavage assay. **c,** Representative denaturing polyacrylamide gel of target cleavage assay using apo-MIWI or MIWI piRISC. A digital overexposure is shown below the gel. **d,** MIWI piRISC loaded with synthetic guides differing at their 5′ or 3′ termini was assayed for the ability to cleave a complementary target RNA. **e,** Northern blot to estimate the yield of purified MIWI piRISC. **f,** Silver stained gel showing purified MIWI piRISC with and without an amino-terminal epitope tag. **g,** Target cleavage assay comparing the activity of piRISC with and without an amino-terminal epitope tag. Data were fit to Equation 1 (see Methods); error of fit (SD) is reported. **h,** Mass spectrometry was used to determine arginine methylation status for recombinant, affinity-purified MIWI (n = 3); arginine methylation status of endogenous MIWI from mouse testis is provided for comparison[77]. Arginine position reports the amino acid number in endogenous mouse MIWI and in parentheses, the corresponding residue in the epitope-tagged recombinant protein. Only modified arginine residues detected in all three independent preparations are reported. N.D., not detected.

Arif et al., Extended Data Figure 1







MIWI piRISC:	Tagged		Untagged	
	Prep 1	Prep 2	Prep 1	Prep 2
$[E]_{active}$ (nM)	0.5	0.1	0.6	0.06
k_{burst} (min^{-1})	0.04	0.03	0.04	0.07

h

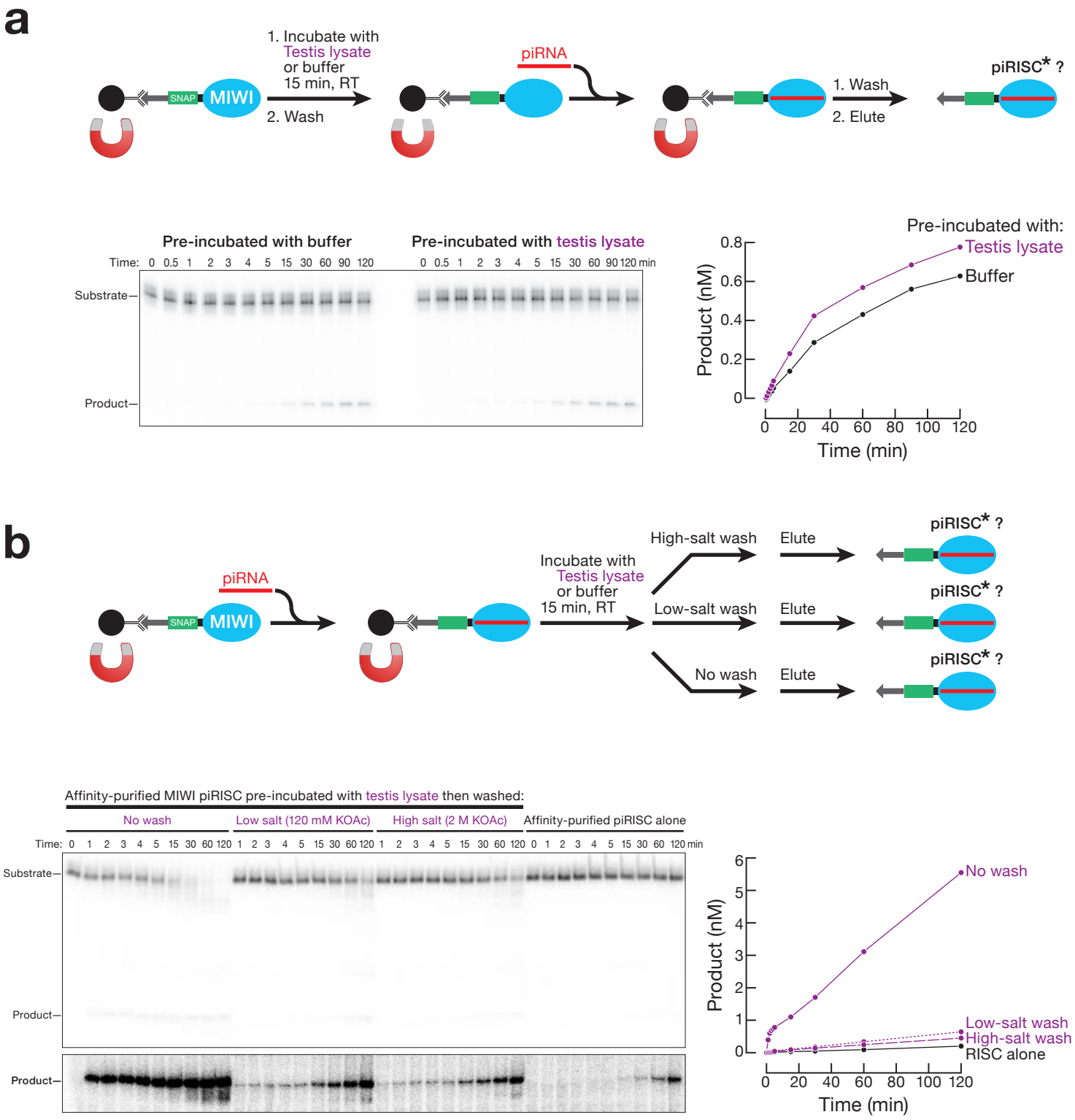
Methylation status by mass spectrometry

Arginine position	Endogenous MIWI Vagin et al., (2009)	Recombinant MIWI
R4 (R244)	n.d.	Mono
R6 (R246)	n.d.	Mono
R14 (R254)	Mono, dimethyl	Mono
R49 (R289)	Mono	Mono
R371 (R611)	Mono	Unmodified
R590 (R830)	Unmodified	Mono

Extended Data Figure 2 | The target-cleavage potentiating component in testis lysate component transiently interacts with MIWI.

a, Immobilized, unloaded apo-MIWI was preincubated with testis lysate, purified as depicted, and then assayed for target cleavage activity. **b,** Immobilized MIWI piRISC was preincubated with testis lysate. purified as depicted, and then assayed for target cleavage activity. A digital overexposure is shown below the gel.

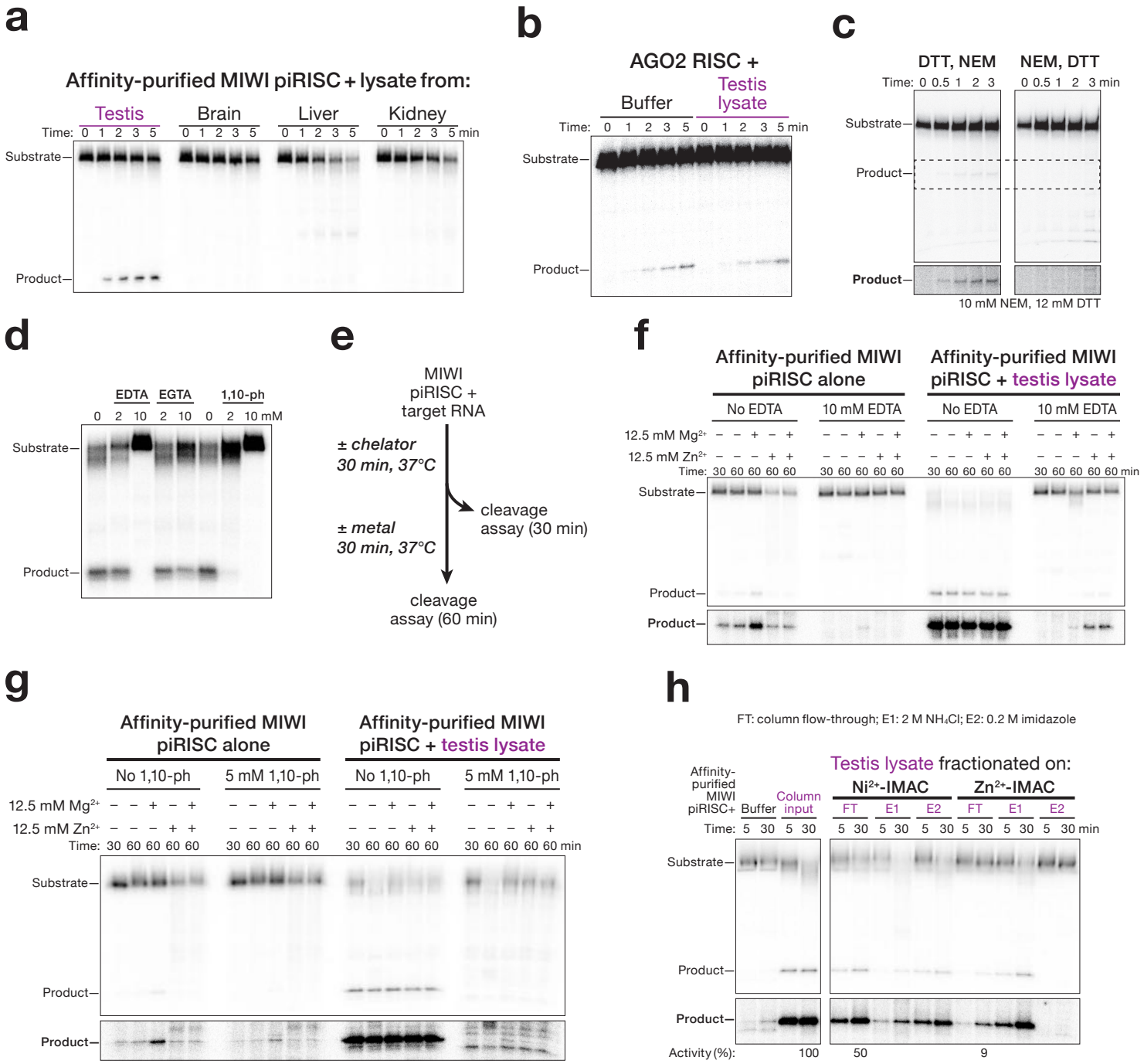
Arif et al., Extended Data Figure 2



Extended Data Figure 3 | Biochemical properties of the MIWI potentiating component in testis lysate.

a, The MIWI-potentiating factor is specific to testis lysate. **b,** Testis lysate does not enhance target cleavage by purified mouse AGO2. **c,** The MIWI-potentiating factor is sensitive to alkylation by *N*-ethylmaleimide. **d,** Effect of divalent cation chelators on MIWI-potentiating activity. For 0 nM chelator, water was added for EDTA and EGTA and ethanol for 1,10-phenanthroline. **e,** Strategy to test metal rescue after incubation of testis lysate with divalent cation chelators. **f, g,** Metal rescue experiments were performed as in (**e**) using EDTA (**f**) or 1,10-phenanthroline (**g**). **h,** Fractionation of testis lysate using Zn^{2+} and Ni^{2+} immobilized metal affinity chromatography (IMAC).
In **c** and **f-h**, a digital overexposure is shown below each gel. Numbers below the gel indicate relative product produced in each lane.

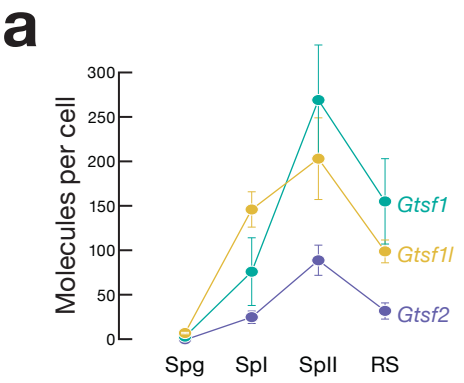
Arif et al., Extended Data Figure 3

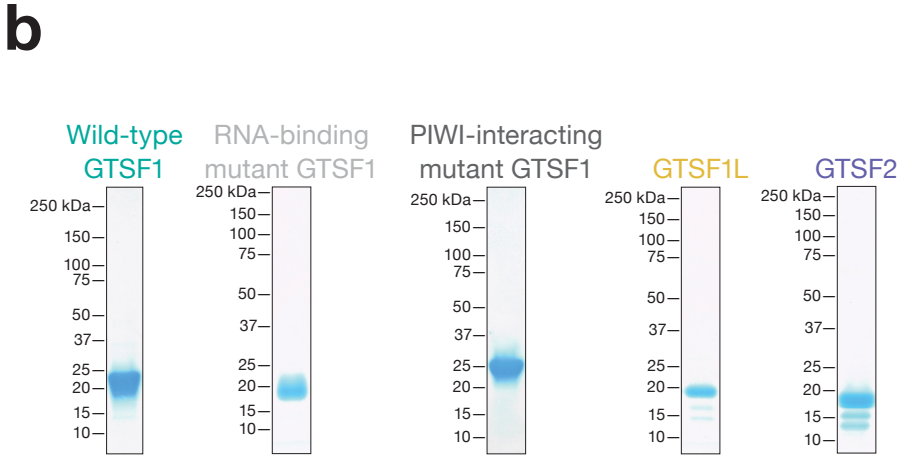


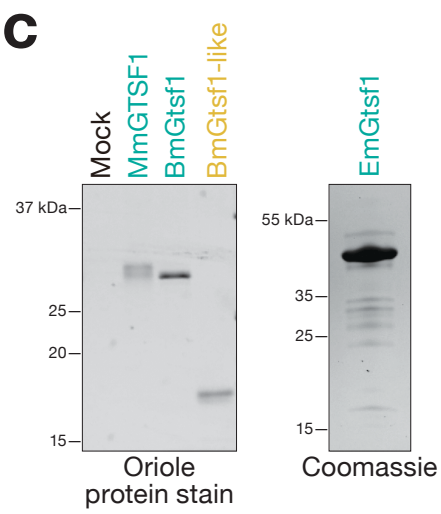
Extended Data Figure 4 | *Gtsf1*, *Gtsf1l*, and *Gtsf2* mRNA abundance and recombinant protein expression.

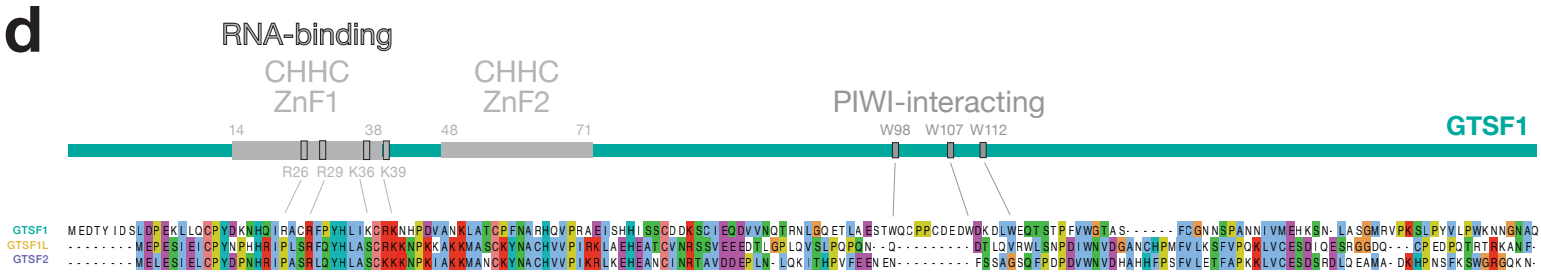
a, For each germ cell type, mRNA abundance (from previously published data; see methods) is reported as mean ± SD (n = 3) in molecules/cell. Spg: spermatogonia; SpI: primary spermatocytes; SpII: secondary spermatocytes; RS: round spermatids. **b,** Coomassie-stained SDS-PAGE of purified, recombinant wild-type GTSF1, GTSF1 mutants, and wild-type GTSF1Like and GTSF2. **c,** Oriole protein stain detection of purified, recombinant, C-terminal V5SBP-tagged GTSF1 orthologs. Mm: *Mus musculus* (mouse); Bm: *Bombyx mori* (silkmoth). Coomassie protein stain detection of purified, recombinant *Ephydatia muelleri* Gtsf1. **d**, Clustal Omega amino acid sequence alignment of the three mouse GTSF1 paralogs.

Arif et al., Extended Data Figure 4









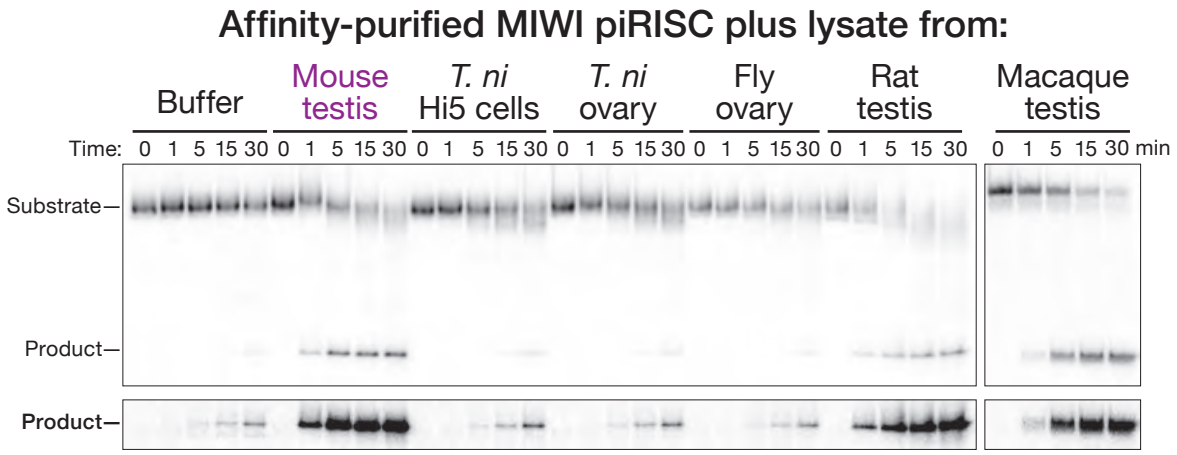
Category	Residue	Threshold, Residue group
Basic	K,R	>60%, KR; >80%, K,R,Q
Acidic	E	>60%, KR; >50%, QE; >85%, E,Q,D
	D	>60%, KR; >85%, K,R,Q; >50%, ED
Polar	N	>50%, N; >85%, N,Y
	Q	>60%, KR; >50%, QE; >85%, Q,E,K,R
	S,T	>60%, WLVIMAFCHP; >50%, TS; >85%, S,T
Unconserved	any/gap	none of the above criteria met

Category	Residue	Threshold, Residue group
Hydrophobic	A,I,L,M,F,W,V	>60%, WLVIMAFCHP
	C	>60%, WLVIMAFCHP
Cysteine	C	>85%, C
Glycine	G	>0%, G
Proline	P	>0%, P
Aromatic	H,Y	>60%, WLVIMAFCHP; >85%, W,Y,A,C,P,Q,F,H,I,L,M,V

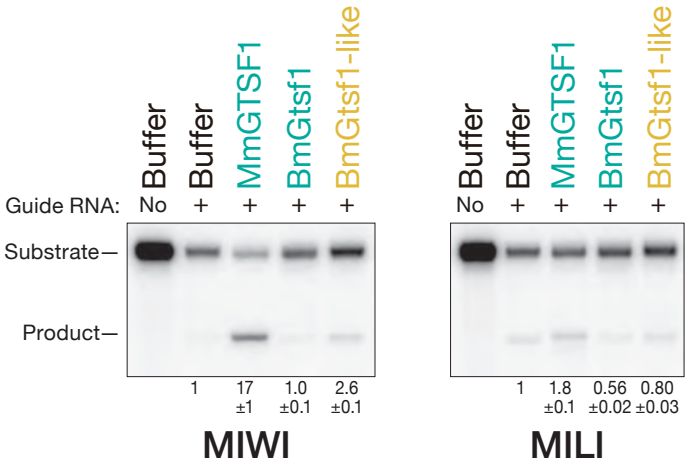
Extended Data Figure 5 | GTSF proteins are PIWI selective.

a, MIWI piRISC target cleavage assay in the presence of lysates from different animals. *T.ni: Trichoplusia ni.* Macaque: rhesus macaque. A digital overexposure is shown below the gel. **b,** Representative target cleavage assay using the indicated PIWI protein in the presence of different GTSF1 orthologs (100 nM). Numbers below the gel lanes report fraction target cleaved (mean ± SD; n = 3). **c,** Left: Percent identity of different GTSF1 orthologs Right: Unrooted tree of GTSF1 orthologs.

a

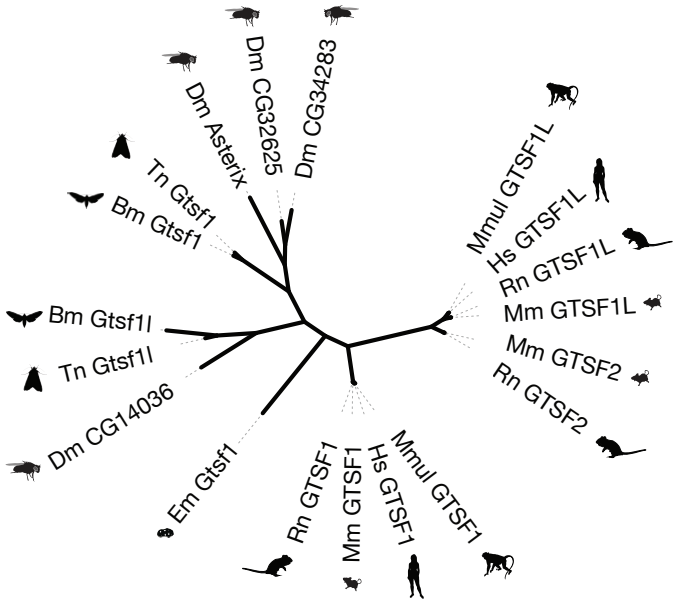


b



c

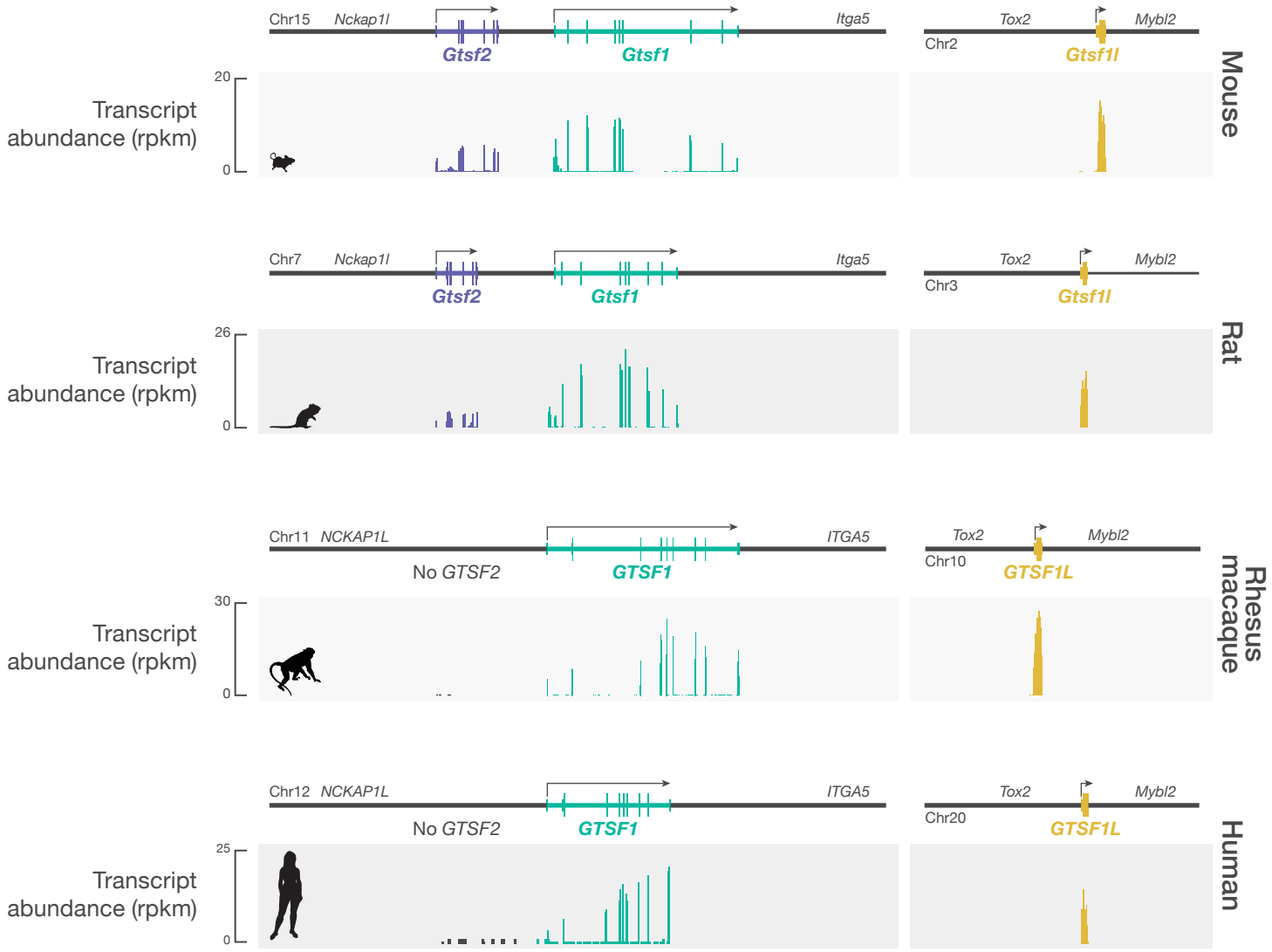
	% Identity (GTSF1)	
Drosophila melanogaster (Dm)	100.00	19.86
Bombyx mori (Bm)	20.00	22.38
Trichoplusia ni (Tn)	19.40	21.17
Mus musculus (Mm)	19.86	100.00
Rattus norvegicus	19.86	100.00
Macaca mulatta (Mmul)	20.55	91.62
Homo sapiens (Hs)	20.55	91.62
Ephydatia muelleri (Em)	14.97	21.47



Extended Data Figure 6 | Expression of GTSF1 paralogs in mouse, rat, rhesus macaque, and human whole testes.

Gtsf genes are syntenic in mice, rats, macaques and humans. Primates, including humans, do not express GTSF2. RNA-seq from total testis RNA of the indicated animals sourced from previously published datasets (see methods).

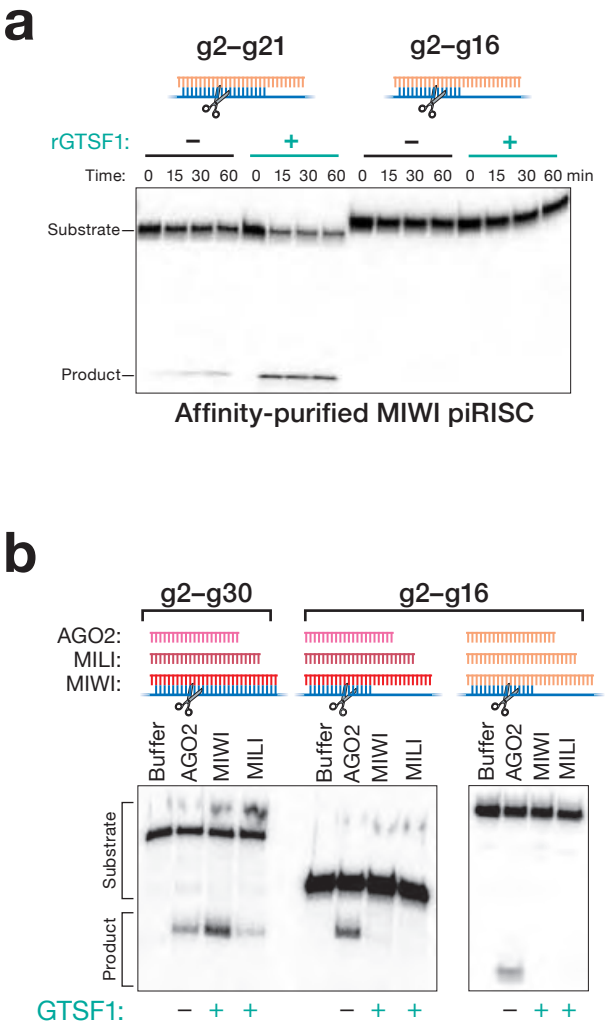
Arif et al., Extended Data Figure 6



Extended Data Figure 7 | Complementarity and GTSF1 requirements for target cleavage by AGO2, MIWI, and MILI.

a, MIWI target cleavage ± GTSF1 using synthetic piRNA guide 1. **b,** Target cleavage by AGO2 siRISC compared with MIWI or MILI piRISC loaded with synthetic piRNA guide 1 or 2, for targets with complete (g2–g30) or partial (g2–g16) complementarity to the piRNA.

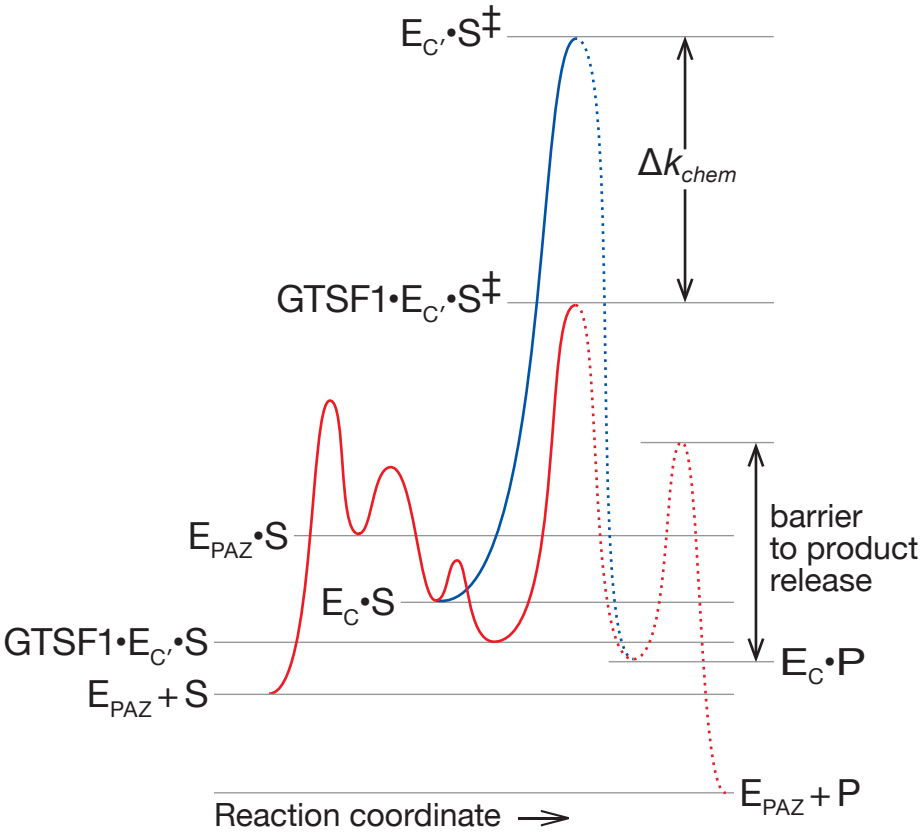
Arif et al., Extended Data Figure 7



Extended Data Figure 8 | Model predicting free-energy change during GTSF1-induced potentiation of target cleavage by PIWI.

Two conformational changes in piRISC (E) are proposed to be required for efficient target cleavage catalyzed by PIWI proteins. E_{PAZ}: piRISC with the 3′ end of the piRNA bound to the PAZ domain; E_C: piRISC with the piRNA fully paired to the target RNA, i.e., the pre-catalytic conformation; $E_{C'}$: E_C in the catalytically competent conformation; $S^{\ddagger}$: transition-state. For piRNAs of biologically relevant length, extensive complementarity is proposed to promote the first conformational change; GTSF1 is proposed to promote the second. Blue: energy barrier to catalysis without GTSF1, i.e., the spontaneous conversion of E_C to $E_{C'}$.

Arif et al., Extended Data Figure 8



Supplementary Table 1. Oligonucleotides and DNA used in this study

Name	Sequence (5′-to-3′)	Notes
AGO/PIWI RNA guides (m, 2′-*O*-methyl; phos, 5′ phosphate)		
p30.47	/phos/UGA GGU AGU AGG UUG UAU AGU AUC CAG AGG	First 21 nucleotides correspond to *let-7* miRNA sequence; 3′,2′ OH; orange label in figures
p30.50	/phos/UGA GGU AGU AGG UUG UAU AGU AUC CAG AGmG	
p30.50DNA	/phos/TGA GGT AGT AGG TTG TAT AGT ATC CAG AGmG	DNA guide
p21.46	/phos/UGA GGU AGU AGG UUG UAU AGU	
p30.53	/phos/UUA GGU AAC CCA GUA GAU CCA GAG GAA UUmC	Guide sequence targeting luciferase in pGL2 plasmid; red in figures
p26.53	/phos/UUA GGU AAC CCA GUA GAU CCA GAG GmA	
p26.53 (3'OH)	/phos/UUA GGU AAC CCA GUA GAU CCA GAG GA	
p21.1175	/phos/UUA GGU AAC CCA GUA GAU CCA	
p16.53	/phos/UUA GGU AAC CCA GUA G	
FLAG-Siwi/MIWI/MILI pre-piRNA	/phos/UCG AAG UAU UCC GCG UAC GUU AUG CUA GCU GAU CCU GAC CGC UGA GUC GU	
AGO passenger strand		
p21.443	/phos/UAU ACA ACC UAC UAC CUC AUU	PS for 21.46
p21.1176	/phos/UGG AUC UAC UGG GUU ACC UUC	PS for 21.1175
Target RNA oligos		
5′ CH_3O/Tar3	/5′ CH_3O/UAU AGA AAA AAU AAC UAU ACA ACC UAC UAC CUC AAU GCU	2-21 matched target against let-7 type guide sequences
5′ CH_3O/Tar4	/5′ CH_3O/UAU GAA GAG AUC GUA ACC ACA ACC UAC UAC CUC AAU GCU ACA	2-16 matched target against let-7 type guide sequences
Primers for IVT targets		
Pr8F_T7_pGl2	GCG TAA TAC GAC TCA CTA TAG GGT TTT AAT GAA TAC GAT TT	Forward Primer to add T7 promoter to template
Pr9R	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTG AGG TAG TAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G	Reverse primer for fully complementary let-7 target
Pr_let7_(2-16)R	GCA CTG TTT TCA TGA AGA CAT GAA GAG ATC TGT AGC ATT GAG GTA GTA GGT TGT GGT TAC GAT CTC TTC ATG AGG AAT TCA TTA TCA GTG	Reverse primer for 2-16 let-7 target
Pr_Luc_Tar1R	TCC GGA ATG ATT TGA TTG CC	Reverse primer for fully complementary luciferase target

Pr_Luc_tar8R	GCG GAA GGG CCA CAC CCT TAG GTA ACC CAG TAG ATC CAC CTC TGG ATA TTA TCA GTG	Reverse primer for 2-21 match luciferase target
Pr_Luc_tar7R	TAT GCG GAA GGG CCA CAC CCT TAG GTA ACC CAG TAG ATC CTC CTC TGG ATA TTA TCA GTG	Reverse primer for 2-20 match luciferase target
Pr_Luc_Tar6R	TTC TAT GCG GAA GGG CCA CAC CCT TAG GTA ACC CAG TAG ATC TTC CTC TGG ATA TTA TCA GTG	Reverse primer for 2-19 match luciferase target
Pr_Luc_Tar5R	GCG GAA GGG CCA CAC CCT TAG GTA ACC CAG TAG ATA TTC CTC TGG ATA TTA TCA GTG	Reverse primer for 2-18 match luciferase target
Pr_Luc_Tar4R	TTC TAT GCG GAA GGG CCA CAC CCT TAG GTA ACC CAG TAG AAA TTC CTC TGG ATA TTA TCA GTG	Reverse primer for 2-17 match luciferase target
Pr_Luc_tar3R	TAT GCG GAA GGG CCA CAC CCT TAG GTA ACC CAG TAG GAA TTC CTC TGG ATA TTA TCA GTG	Reverse primer for 2-16 match luciferase target

Sequence-affinity purification (m, 2′-*O*-methyl)

luciferase capture oligo	/Biotin/mAmCmU mAmCmU mGmCmA mGmCmA mCmAmA mCmCmC mUmAmC mCmAmA mAmUmU mAmCmC mUmAmA mCmUmG mC	
luciferase competitor oligo	GCA GTT AGG TAA TTT GGT AGG GTT GTG CTG CAG TAG T	DNA oligo
let-7 capture oligo	/Biotin/mAmUmA mGmAmC mUmGmC mGmAmC mAmAmU mAmGmC mCmUmA mCmCmU mCmCmG mAmAmC mG	
let-7 competitor oligo	CGT TCG GAG GTA GGC TAT TGT CGC AGT CTA T	DNA oligo

Primers for BmAgo3 targets

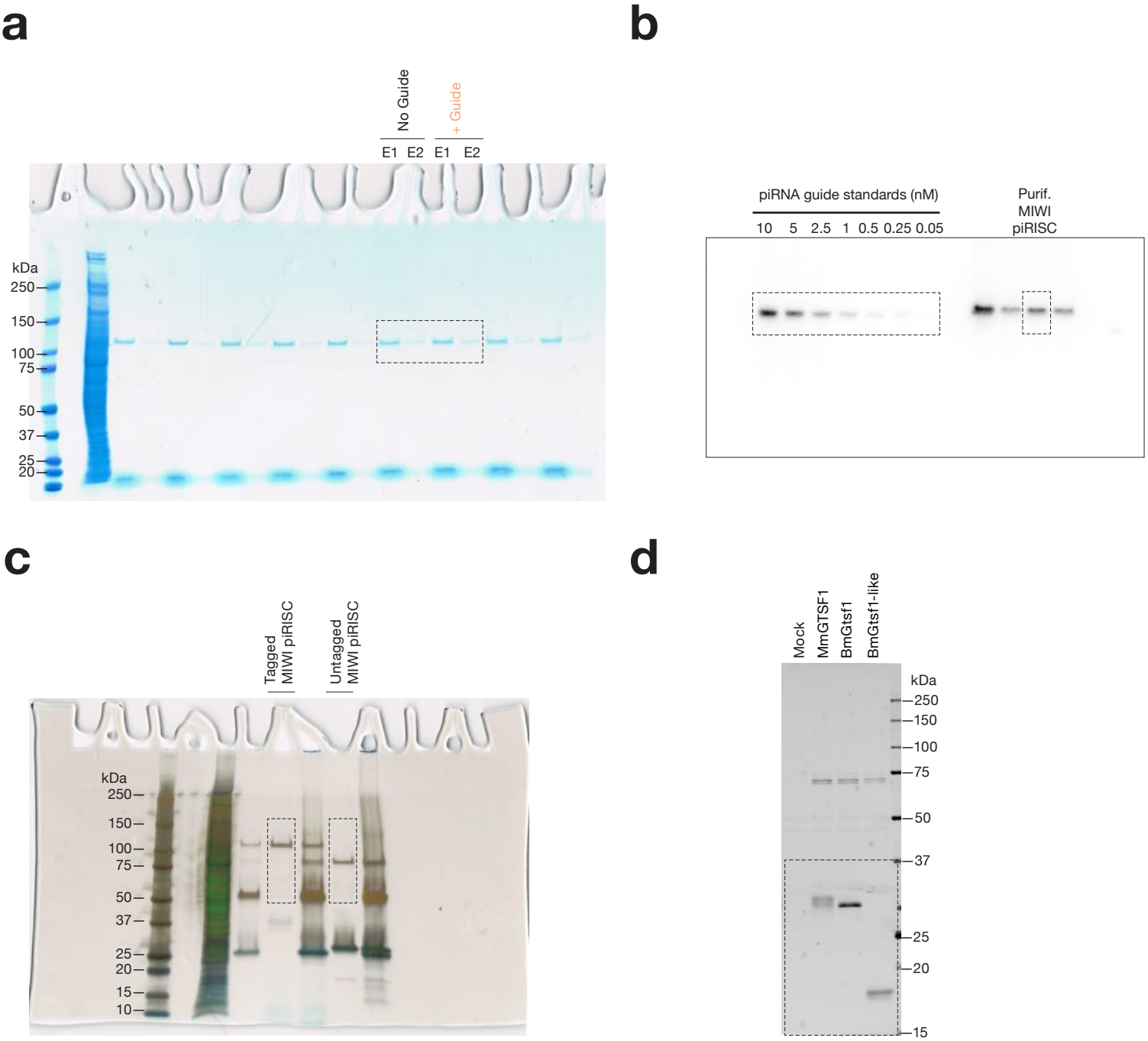
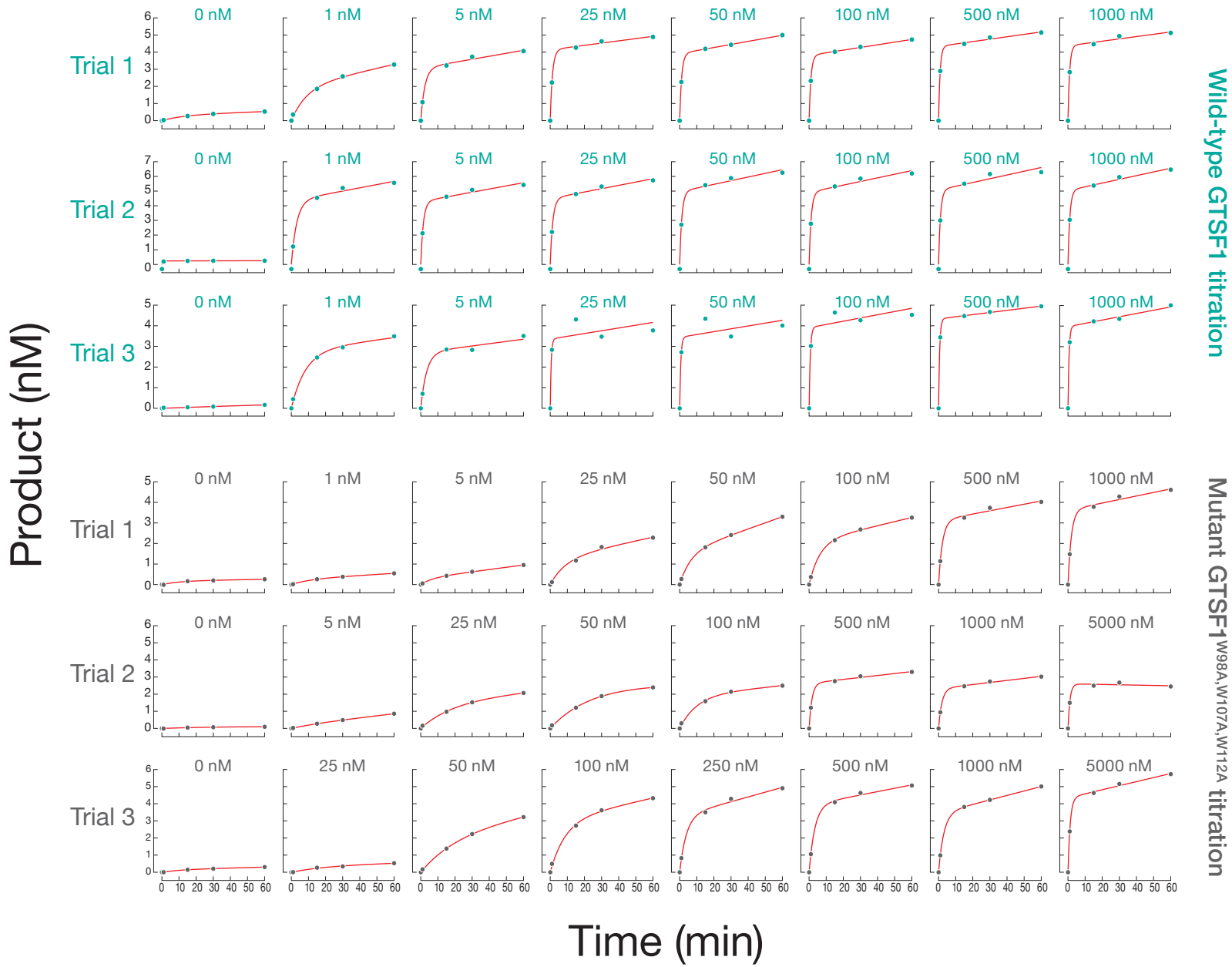
Target-F	GCG TAA TAC GAC TCA CTA TAG TCA CAT CTC ATC TAC CTC C	Forward Primer to add T7 promoter to the template
Target #1 (pi848-perfect target-R)	AGG TGA CAC TAT AGA TTT ACA TCG CGT TGA GTG TAG AAC GGT TGT ATA AAA GGT CGA GTG ACA TGT TTG GAA GCA ACA GCA TCC AGA GGA GTT CAT G	Reverse primer for fully complementary target
Target #2 (pi2148ds-perfect target-R)	AGG TGA CAC TAT AGA TTT ACA TCG CGT TGA GTG TAG AAC GGT TGT ATA AAA GGT AGG GCC AAG ACC AAA TAT AGT ACG TTT CAT CCA GAG GAG TTC ATG	Reverse primer for fully complementary target

Construction of *Gtsf1*$^{KI/KI}$ mouse

HDR donor	ACTTTGTCCAATTCACCGACAGCCTGTTATTGGGGGAAAGGAATCAA ATCACCTGTTAACAATTTGATGGGGAAGATAAGAGACAGCTGAATGA TGGTGTTATGAGTTCCAGCGTCCTTGGCCTACTTGGAATCAGTCTGT TTGCTAAGAAAGAATAGGTCAGCTCTCTAAGTCGCTGAATATCTTTA CTCATCTTTTTCAAAATAATAATAATAATGTAATTCTTGGTTTTTCCTT AAAACCTTTTTTTATTTTTAACAAAGTATGACATAACTGTAGTAGAAGT TCCATACGGGTTAAGATAGTTTAACTAATTCTCAGATAGTTTCCCCAA TGATTGCCATTAACAAATGGGCTCTGGAGCAAAGTCCGAAACTAGAG TTTAGTTGATCACCCCTATTGCTTAACAAAGGTTCTAATGCTGCCTAT GTCATAGAATGGGGGGGAAGGGGGTGTGGAATTTTTGTGAGAGGC CAAAGGAGAGCTTACCTAGAGACAACAGCTGTATGTGCATCCAATTT TCAGAAGATTCCTATGGGAGAGGTTATATACAAAATTAGTTTTCCCAT AATACGTGAACTAAATCAGTGTTCAATAGAGATTCCAGTAAAAGAGA ATTCAAACTCAGAGACATTTCCTCTTTTAATTTTTTTCATGTTTGTTAT ATTTGTTCTCTTAGATGTTCTCTATGAGCTCTTAATTTTTTTAACATTC ACATATATTCCCTCAGTGCAGTTAAGTAATCAGTACTGACTAAAACAG TATTTATCTTGTTATTATATAGAAATGAGAGGAACAAAATGAAAAATG GCTAGAAAGAGAGAATATTGGTGTGACGTTTTCTGATCATGATTAAG AAAAGTTACTAAGCTGGGTATGGTTGAAAATGCCTTTAATCCCAGAA CTCAGGAGGCAGAGGCAGGTGGATCTCTGTGTGTTCAAGGTCAGTC TGGTCTGCATATTGAGTTCCAGGACAGCCAAAACTATACAGAGAAAC CTTGTCTCTTAAAAAAGAAGAAAAAGAAAGGTCATTAAATGAGATGTA GCTTTTATTTTACATGTGTTTGTGTGCATTTTGCCCACAGAAGCCAGA AGAGGGCATCTGATCCCCTGAAACTGGAGTTATAGGCAGTTGTAAG CCATGTGGGTGCTGAGAATCAAACCCAGATGTCACTAAGAACACCC AGAGCTCTTAACCATTTAATCATCTTTCCATCTGGAAAGAAGATCTTT AAAGTAGTTTATCCAGAAGTCTAAATACTGGTATTTTTATGACCTTGA CTTGGCCCCCATTACTTATGTTAAATACCATAACCCCTAATAGCATGT AAACAGTGGCAGGTGTTGTAATACCACTTAAAATAAATTGTTATTATT GTCACCTGAATGTAGTTTGTCAGTTGTCATATTTGACCTAGCAATTTA TTCCAAGAATGTGAACTATGCAGTGATGCTAAACTTAGTCATAGGAG TAAAGAATGCCTTTAAGAAGAAACTGGAGAGAAGGTTGTATGCTAAG GGTTTATTTTATTGAACTTGGGGCAGCTTCAGGTCCAGGAATGACTC CACATCTTAGGGCTACTATTTGATGCGTTTCATTTCTGATTCTCTGAA CCATGCTTTTTCCTCTAGCACTTGAACAGGACTCCTTTGCCATCAGC **ATG**GACTACAAGGACCACGACGGTGACTACAAGGACCACGACATCG ACTACAAGGACGACGACGACAAGGGCTCCGGATACCCATACGATGT GCCAGATTACGCTGGATCAGGCGAAGACACTTACAGTATGTACTACT TGCTGCTTTTAGAAAATAATACAATATTTATGTGTAGTTTGATGCCTG ACTAGACTGTATGTTGTGAAAACTTCAGATTCCTCCCTGAAATGCTG AGAAAGAAAGGTCTGTGTGATATTGAGGTATACAGATGCTCAAATTT TAAATAAGTAACTTTAGTCTACAGCCATATTACCCTGAAAAAAAATGG TCTGAAACAAGCACTTTAAACTTAAGTAATTAGACATATGAGAACCAA CTGTAATGATACACAACTGTAATCCCAACTCATAGGAAGCTGAATCA TGGGGATCACAACTTCTAGGCCATCATGGGCTTCACAGAAAGACCA TCAAAAGATAAACCAGTAGTCCACTGAGATGGCTCAAGGGTGAAGT GCTTATTACCATAAAAGCCTGACAACCTTAGTTTGAATTTTAGAACTC TTGTAAAAGTAGGAAAGAAAACTGAATCTATAAAGTCGTCCTCTGAC CTACACACAGACCACACACAATAATAAGTACAAAATTTTTACAATACT GGGGATGTAACTCAGTGGTAGAGCATTTGTGTAGCATTCTCAAGACC CATATTTTAATACCAAGATAATAGGTAGAGAGAGAGAGAGAGATATG AATGATTGATTGGTGGTTAAATTGGTGACCAAGATAGATGGATGAAT GGATGATTGATAAGTAGATGGATGGGTGGATGGATGTAGATAAATAG ATGGATCATTGATAGATAGATAGATAGACAGATTCTTTTCAAGGTAGA ACAGTAAGTTAGAGCTGACCTAAATTGCCAGACTAAGTAAAGCATTT ATGGTAGTTCTCATTATCAACTTTGACACATGCTCTATATGAGCAGTT ACAATATGTGTGAATCATGGGGTACCTCTGTTCTGTCAACTGCTGAC ATTTGCTAGTGATAGAATGGATTAAATGAAGGAAAACTAATAAGAATT TCATCTGTGACTGCTCCCTGGGTGTTCTGAAGCTTCAGAAATAAGAA GAGATTGCTGGCTAAGAGCTCAGTGGCACAGTAGCATGCCTGTCCT CTGGGGCCTTGCCTCCCTGTTCATTCTGTCCTTCAGATCACCTTTCA AGCTGATTTGCATCATGACCGTGGTTAGCAATTTATTCCATGTTTGC CACTCTTCCAGCGAAACTGTCAGTTTCAATTCCCCTCTGGTTTCTGA TTTTCATTGCAGTCTCCTGGTTTATAAGCTTCTTATCAATGGTGAATT GTTTGTCTTGCTTTTATCCTTACCTTCTGGGGTGGCAAGAGTGTTTTA TTAGGTGCCTTTTGCTCTCAGAATGCACTATTTAACTTTCTTTCCTTA ACCAGGCTTGTTAAGATTTGGGGCATAGCCTGAGGGCAGCTTTGAA TAACATGGTATAGGTTCTGGGTTTTCTTGACACAGCAGTCTGTGTCT ACAATTGTCTTAGCTGTAGGAGGTTATTTTTTATTTTAGCTGGCCAGT TTGGTAGTCCTTAGATTAAGGAACTTTTTTCTCATCTATGTTCCAAAA CATGAAAAGTAAAATCACCTACCTGACAAAAGTAATGAGATCATTGT CAGGGGAACGAGAGCAGCCAATACAGCATACAGGAAGGAGTAAAG CAAAGAACAAATTGGGGGGAAACTCTAGTTATTAATACTTCATAGAA CATTCAATTGAAGCTATCAACTTCTTAACTAGAAATAGACAAGGATGC TAATGAGATTATTCTCAACTACGGCTAGCTAGATTGAGAGAGTGTTG GAAAGTGCAGTGGGC	Left and right homology arms; **ATG**: *Gtsf1* start codon; 3XFLAG-HA knock-in cassette

sgRNA target sequence	GTCTTCCATGCTGATGGCAAAGG	PAM sequence
Primer F1	CTGTTTTACTGAGACCTGCAACTTT	Genotyping and validation primers
Primer R1	GATCCAGCGTAATCTGGCACAT	
Primer F2	ACTACAAGGACCACGACGGTG	
Primer R2	GACTGGATTGCTGGTGTAACAATA	

Arif et al., Supplementary Data Fig. 1



Arif et al., Supplementary Data Fig. 2