



HAL
open science

Diversification of small RNA amplification mechanisms for targeting transposon-related sequences in ciliates

Masatoshi Mutazono, Tomoko Noto, Kazufumi Mochizuki

► **To cite this version:**

Masatoshi Mutazono, Tomoko Noto, Kazufumi Mochizuki. Diversification of small RNA amplification mechanisms for targeting transposon-related sequences in ciliates. Proceedings of the National Academy of Sciences of the United States of America, 2019, 116 (29), pp.14639-14644. 10.1073/pnas.1903491116 . hal-02387371

HAL Id: hal-02387371

<https://hal.science/hal-02387371>

Submitted on 7 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Diversification of small RNA amplification mechanisms for targeting transposon-related sequences in ciliates

Masatoshi Mutazono, Tomoko Noto and Kazufumi Mochizuki*

Institute of Human Genetics (IGH), CNRS, University of Montpellier, Montpellier, France

*Correspondence and Lead Contact: kazufumi.mochizuki@igh.cnrs.fr

BIOLOGICAL SCIENCES: Developmental Biology

Key words: small RNA, transposon silencing, DNA elimination, ciliate

1 **Abstract**

2

3 The silencing of repetitive transposable elements (TEs) is ensured by signal amplification
4 of the initial small RNA trigger, which occurs at distinct steps of TE silencing in different
5 eukaryotes. How such a variety of secondary small RNA biogenesis mechanisms has
6 evolved has not been thoroughly elucidated. Ciliated protozoa perform small RNA-
7 directed programmed DNA elimination of thousands of TE-related internal eliminated
8 sequences (IESs) in the newly developed somatic nucleus. In the ciliate *Paramecium*,
9 secondary small RNAs are produced after the excision of IESs. In this study, we show that
10 in another ciliate, *Tetrahymena*, secondary small RNAs accumulate at least a few hours
11 before their derived IESs are excised. We also demonstrate that DNA excision is
12 dispensable for their biogenesis in this ciliate. Therefore, unlike in *Paramecium*, small
13 RNA amplification occurs prior to IES excision in *Tetrahymena*. This study reveals the
14 remarkable diversity of secondary small RNA biogenesis mechanisms, even among
15 ciliates showing similar DNA elimination processes, and thus raises the possibility that
16 the evolution of TE-targeting small RNA amplification can be traced by investigating the
17 DNA elimination mechanisms of ciliates.

1 **Significance statement**

2

3 Secondary amplification of small RNA triggers is necessary to silence repetitive
4 transposable elements (TEs) and thus to maintain genome integrity. Despite their
5 importance, these secondary small RNAs are produced by various mechanisms in
6 different eukaryotes. How such a variety of mechanisms has evolved has not been
7 thoroughly elucidated. In this report, we show that in the ciliate *Tetrahymena* secondary
8 small RNA production occurs from the eliminated sequences prior to their excision
9 during the small RNA-directed programmed DNA elimination of TE-related sequences.
10 This situation contrast with the excision-dependent secondary small RNA production
11 observed in another ciliate, *Paramecium*. This study therefore shows the possibility of
12 tracing the evolution of TE-targeting small RNA amplification by investigating DNA
13 elimination mechanisms of ciliates.

14 \body

1 **Introduction**

2

3 Signal amplification of the initial small RNA trigger is important to ensure the silencing
4 of repetitive transposable elements (TEs). Curiously, secondary small RNA biogenesis
5 occurs by various mechanisms that are coupled with distinct steps of TE silencing in
6 different eukaryotes, such as nucleolytic cleavage of TE transcripts, recruitment of RNA-
7 dependent RNA polymerase, and heterochromatin-directed transcription (1–5). To
8 understand how such a variety of small RNA amplification mechanisms have evolved, it
9 is important to compare evolutionarily closely related TE-silencing processes that have
10 distinct secondary small RNA biogenesis mechanisms. This study raises the possibility
11 that the programmed DNA elimination mechanisms of ciliated protozoa can be used for
12 such a comparative evolutionary biology approach.

13

14 Like most other ciliates, each cell of *Tetrahymena thermophila* (hereafter referred to as
15 *Tetrahymena*) contains two types of nuclei: the transcriptionally inactive germline
16 micronucleus (MIC) and the transcriptionally active somatic macronucleus (MAC).
17 During the sexual process of conjugation, the MIC produces new MICs and MACs, while
18 the parental MAC is discarded. The MIC genome includes ~12,000 IESs, which are
19 removed from the newly formed MAC. IESs contain many potentially active TEs and their
20 remnants (6). Each IES can be excised in linear or circular form. Because excised IESs in
21 linear form are very short lived (7, 8) while circularized IESs are stable for more than
22 several hours (9), circularized IESs should be produced either directly during the excision
23 process or from linear IESs right after their excision.

24

25 Small RNAs play a key role in the DNA elimination in *Tetrahymena* (10–12). At early
26 conjugation stages (~2-3 hours post mixing [hpm]), ~60% of IESs, called Type-A IESs,
27 and their genomic surroundings are bi-directionally transcribed in the MIC (13, 14), and
28 the transcripts are processed to ~26-32-nt primary small RNAs called early-scnRNAs (15,
29 16). Early-scnRNAs are loaded into the Piwi-clade Argonaute protein Twi1p (17, 18) and
30 move into the parental MAC, where those complementary to the MAC genome (i.e., non-
31 IES sequences) are degraded by the process called scnRNA selection (14, 19). The
32 remaining IES-specific early-scnRNAs are then shuttled into the developing new MACs
33 (~7-8 hpm) and are believed to base pair with nascent transcripts from Type-A and Type-

1 B IESs, the latter of which constitute ~40% of all IESs and share repetitive sequences with
2 Type-A IESs (20). These interactions recruit the Enhancer of zeste [E(z)]-like
3 methyltransferase Ezl1p, which mediates the accumulation of tri-methylations of histone
4 H3 at lysine 9 (H3K9me3) and lysine 27 (H3K27me3) and the HP1-like protein Pdd1p to
5 IESs (21, 22). This heterochromatin nucleation induces the production of ~26-32-nt
6 secondary small RNAs, called late-scnRNAs (20), and the excision of IESs by the
7 domesticated piggyBac transposase Tpb2p (23, 24). The flanking sequences are re-
8 ligated by non-homologous end joining (NHEJ) (25). Besides Type-A and Type-B IESs,
9 there are ~20 IESs in the genome that produce few scnRNAs and are removed by the
10 other piggyBac transposases, Tpb1p and Tpb6p (26, 27).

11

12 The other ciliate *Paramecium tetraurelia* (hereafter referred to as *Paramecium*) exhibits
13 a very similar DNA elimination process, including Dicer-dependent scnRNA production
14 (28); loading of the small RNAs into Piwi-clade Argonaute proteins (29); scnRNA
15 selection (28, 30); scnRNA- and an E(z)-like protein-dependent accumulation of
16 H3K9me3 and H3K27me3 on IESs (31); IES excision by a piggyBac transposase (32, 33);
17 and repair of the excision sites by NHEJ (34). Recently, it has been reported that the
18 secondary small RNAs, called iesRNAs in *Paramecium*, are produced after the excision of
19 IESs (35). The excised IESs are concatenated and circularized and proposed to serve as
20 templates for the production of precursors for iesRNAs in this ciliate (35, 36). On the
21 other hand, although heterochromatin formation has been shown to be essential for the
22 biogenesis of late-scnRNAs in *Tetrahymena* (20), it remains to be determined if the late-
23 scnRNA production is dependent on IES excision. To examine whether the post-excision
24 production of secondary small RNAs is conserved, we investigated the causality between
25 the excision of IESs and the biogenesis of secondary small RNAs in *Tetrahymena*.

1 **Results**

2

3 **Highly sensitive assay for determining the timing of de novo IES excision**

4 To understand the temporal relationship between the excision of IESs and the production
5 of late-scenRNAs in *Tetrahymena*, we first aimed to determine the timing of IES excision.
6 It has been reported that de novo IES excision can be detected by PCR amplification of the
7 ligation junctions of excised IES circles (jPCR) (9, 37). However, the sensitivity of their
8 detection has not been examined. To detect de novo DNA elimination events with high
9 sensitivity, we have established a method combining jPCR with the digestion of linear
10 DNA (DLD) and rolling circular amplification (RCA), referred to as DLR-jPCR (Figure 1A).

11

12 Two wild-type *Tetrahymena* cultures possessing different mating types were mixed to
13 initiate conjugation, and total genomic DNA was extracted at the onset of mixing and
14 subsequently at 2-hr intervals from 4 to 20 hr post mixing (hpm). The genomic DNA was
15 then used in DLR-jPCR assays. We chose 9 representative IESs, including 3 each of the
16 major IES types, Type-A, Type-B and *TPB1/6*-dependent IESs, to investigate the
17 appearances of IES circles. As shown in Figure 1B, IES circles were specifically detected
18 in the late stages of conjugation (Figure 1B, RCA+), consistent with the fact that DNA
19 elimination occurs during the development of the new MAC, which is formed at ~7 hpm.
20 IES circles were detected for all 9 tested IES loci, indicating that DLR-jPCR can detect de
21 novo IES excision events, regardless of the type of IES.

22

23 The timing of the appearance of IES circles varied among different IESs. For example,
24 circles of IES737 and IES1147 were detected at 10 hpm, whereas those of the other IESs
25 only appeared after 12 hpm (Figure 1B). Although we have failed to find any obvious rule
26 defining the timing of DNA excision for different IESs to date, these results are consistent
27 with the previous notion that some IESs are eliminated earlier than others (7). IES circles
28 were detected continuously for at least several hours. A previous report (9) suggested
29 that this situation is caused, at least in part, by the persistence of each IES circle for
30 several hours before it is degraded, although cell-to-cell variability of the timing of IES
31 excision may also contribute to the continuous detection of IES circles.

32

1 We failed to detect any IES circles when we performed similar reactions without RCA
2 (Figure 1B, RCA-), indicating that the RCA step significantly sensitizes the detection of IES
3 circles. To further assess the sensitivity of DLR-jPCR, we prepared serial dilutions of
4 genomic DNAs after the RCA reaction and performed jPCR. As shown in Figure 1C, we
5 found that DNA circles from IES737 could be detected even when genomic DNA from 12
6 hpm was diluted 2,000 times. This result indicates that DLR-jPCR is probably >2,000
7 times more sensitive than the previous IES circles detection methods, which do not
8 include the RCA step, and the dynamic range of the detection sensitivity of the DLR-jPCR
9 assay spans more than three orders of magnitude.

11 **Late-scnRNA production is initiated before IES excision becomes detectable**

12 We next sought to compare the timing of the appearance of IES circles with that of late-
13 scnRNAs. When late-scnRNAs are produced in the new MAC, early-scnRNAs expressed
14 from the MIC at earlier stages of conjugation persist. Early- and late-scnRNAs cannot be
15 distinguished by simple biochemical properties, such as length and base composition
16 (20). Therefore, to specifically investigate late-scnRNAs, we focused on Type-B IES-
17 derived small RNAs, which are predominantly, if not exclusively, late-scnRNAs (20).

19 Small RNAs extracted from mating wild-type cells at 3, 6, 8, 10 or 12 hpm were sequenced
20 (small RNA-Seq), and those that uniquely mapped to the three studied Type-B IESs
21 indicated above were counted. As shown in Figure 2A-C, small RNAs derived from these
22 Type-B IESs remained at a low level till 6 hpm and clearly increased after 8 hpm,
23 indicating that late-scnRNAs began to be produced from these IESs at 8 hpm at the latest.
24 A similar result was obtained when small RNAs were mapped to each of the 3,293 Type-
25 B IESs (Figure 2D). The earliest time point at which IES circles were detected for these
26 IESs was 10 hpm (IES737) or later (IES5, IES4092) (Figure 1B). Therefore, late-scnRNAs
27 expressed from an IES can be detected by small RNA-Seq before the excision of that IES
28 becomes detectable by DLR-jPCR.

30 Although these results most likely indicate that the production of late-scnRNAs occurs
31 prior to the excision of IESs, the above data do not exclude the possibility that late-
32 scnRNAs are produced after IES excision, but the detection of late-scnRNAs by small RNA-
33 Seq is more sensitive than that of IES circles by DLR-jPCR. Therefore, we next aimed to

1 use a genetic mutant that is defective in IES excision to further investigate the
2 relationship between the production of late-scRNA and the IES excision.

3

4 ***TPB2*-null mutants block the accumulation of the Tpb2p excisase**

5 Tpb2p is important for the DNA elimination of Type-A and Type-B IESs in vivo (23) and
6 catalyzes DNA double-strand breaks to form a 4-base 5' overhang in vitro (24), which is
7 identical to the intermediate structure observed in DNA elimination in vivo (9). Therefore,
8 Tpb2p is believed to be the enzyme catalyzing the final excision step of Type-A and Type-
9 B IESs. Previous loss-of-function studies of *TPB2* were conducted using RNAi (23) or
10 conditional (24) knockdown. In these studies, IES elimination was not completely
11 blocked, probably due to incomplete loss of *TPB2* expression. Therefore, the conditions
12 applied in these previous genetic studies cannot be used to explicitly determine the
13 requirement for *TPB2* in late-scRNA biogenesis.

14

15 To produce *TPB2*-null mutants, we employed CRISPR-based mutagenesis technology.
16 Although CRISPR technology was used for mutagenesis of *Tetrahymena* in two previous
17 reports (38, 39), we have made several modifications to the previous protocols to
18 efficiently obtain *Tetrahymena* mutants (see SI Materials and Methods). We introduced a
19 construct expressing Cas9 and gRNAs into the MAC of vegetative cells of two wild-type
20 strains (B2086 and CU428) with different mating types. We expected the Cas9-gRNA
21 complex to introduce a mutation in the MIC (Figure 3A). We used two gRNAs that target
22 slightly different positions of the first exon of *TPB2* (Figure 3B). After the induction of
23 Cas9 expression, the cells were cloned, and the *TPB2* locus in the MIC was amplified by
24 PCR using primers “a” and “c” shown in Figure 3A. The products were then sequenced.
25 The PCR amplicon contains a chromosome breakage sequence (Cbs) that remains intact
26 in the MIC but is severed to form the ends of two separate chromosomes in the MAC
27 (Figure 3A). This PCR assay therefore specifically amplifies the *TPB2* MIC locus.
28 Approximately half of the clones examined contained insertions or deletions around the
29 targeted region. We choose two clones each from the B2086 and CU428 strains that
30 contained one of the three frameshift mutations in *TPB2*, which were referred to as *tpb2-*
31 *1fs*, *tpb2-2fs* and *tpb2-3fs* (Figure 3B). We did not detect any copies of wild-type *TPB2* in
32 the MIC in these cell lines, indicating that both copies of *TPB2* in the diploid MIC were
33 disrupted in these *TPB2* mutants.

1

2 The *TPB2* MAC locus of the *TPB2* mutants was also analyzed by PCR using primers “b”
3 and “c” shown in Figure 3A. Although this PCR assay was designed to amplify both the
4 MAC and MIC loci, the MIC locus was underrepresented because of the difference in
5 ploidy between the MAC (~50n) and the MIC (2n). We detected only the wild-type *TPB2*
6 locus in these PCR products, indicating that the CRISPR system used here did not
7 efficiently induce mutations in the MAC (Figure 3B). This result might have occurred
8 because Cas9 was fused to part of the MIC-localized protein MLH1 in our experiment to
9 enhance the MIC localization of Cas9 and/or because DNA double-strand breaks might be
10 repaired predominantly by homologous recombination with the sister copies in the
11 polyploid MAC. Nevertheless, Tpb2p detected by western blot analysis was reduced to
12 background levels in the *TPB2* mutants (Figure 3C), indicating that *TPB2* was exclusively
13 expressed zygotically from the new MAC in the wild-type cells. We concluded that the
14 mutants were *TPB2*-null mutants.

15

16 ***TPB2* mutants block DNA elimination of Type-A and Type-B IESs**

17 H3K9me3 and H3K27me3, scnRNA-dependent heterochromatic modifications,
18 specifically occur on IESs in wild-type cells (21). Consistent with a previous report
19 describing RNAi knockdown of *TPB2* (23), these modifications accumulated in the new
20 MAC in the *TPB2* mutants at 8 hpm (SI Appendix, Figure S1). In the wild-type cells at 12
21 hpm, these heterochromatic marks accumulated in nuclear foci referred to as
22 heterochromatin bodies (HBs). The formation of HBs is believed to be triggered by
23 Tpb2p-induced DNA double-strand breaks (40). Consistently, HBs were not detected in
24 the *TPB2* mutants at 12 hpm (SI Appendix, Figure S1). The presence of H3K9/K27 me3
25 and the absence of HBs in the *TPB2* mutants therefore indicate that heterochromatin is
26 formed, but IES excision is generally blocked in these mutants.

27

28 We next directly investigated DNA elimination in the *TPB2* mutants by DLR-jPCR. As
29 shown in Figure 4, we could not detect IES circles from any of the tested Type-A and Type-
30 B IESs at 12 and 18 hpm in the absence of Tpb2p (Figure 4A, *TPB2* mut), whereas IES
31 circles from all of these IESs were detected at one or both of those time points in wild-
32 type cells (Figure 4A, WT). In contrast, IES circles from the three tested *TPB1/6*-
33 dependent IES loci were detected in the *TPB2* mutants, although their appearances were

1 delayed as compared to those in wild-type cells (Figure 4A), Therefore, DNA elimination
2 of all studied Type-A and Type-B IESs was reduced to an undetectable level in the *TPB2*
3 mutants. According to the results of sensitivity testing of DLR-jPCR shown in Figure 1C,
4 we estimate that the frequency of the excision of these IESs in the *TPB2* mutants was
5 reduced to less than $\sim 1/2,000$ of that in the wild-type cells.

6 **Blocking DNA elimination does not result in loss of late-scnRNAs**

7 Finally, to investigate the accumulation of late-scnRNAs in the absence of IES excision, the
8 *TPB2* mutants were mated, and their 26- to 32-nt small RNAs were analyzed. Total
9 amount of small RNAs detected by denaturing gel electrophoresis was similar between
10 the wild-type cells and the *TPB2* mutants at both early- (3 hpm) and late- (10 and 12 hpm)
11 conjugation stages (SI Appendix, Figure S2A). Small RNAs that uniquely mapped to the
12 3,293 Type-B IESs were increased between 3 and 10 hpm in the *TPB2* mutants (Figure
13 4B, top) and small RNAs derived from the three representative Type-B IESs were
14 detected in the *TPB2* mutants at 10 hpm (Figure 4B, bottom) and 12 hpm. In contrast, the
15 levels of Type-B IES-mapping small RNAs remained low at 10 and 12 hpm in *PDD1* KO
16 cells, which lacked the core heterochromatin component (41) (Figure 4B). Because early-
17 scnRNAs are produced only at early conjugation stage (~ 3 hpm) (20), any increase of
18 small RNAs at later stages should be attributed to the production of late-scnRNAs.
19 Therefore, the results above indicate that IES excision, but not heterochromatin
20 formation, is dispensable for the production of late-scnRNAs.

21
22
23 The amounts of Type-B IES-mapping small RNAs at 10 and 12 hpm in the *TPB2* mutants
24 were lower compared to those in wild-type cells (Figure 4B, top). In contrast, Type-A IES-
25 mapping small RNAs were not obviously affected in the *TPB2* mutants (SI Appendix,
26 Figure S2B). It is highly likely that the reduction of Type-B IES-mapping small RNAs, as
27 well as the delayed elimination of the *TPB1/6*-dependent IESs (Figure 4A) in the *TPB2*
28 mutants is indirectly caused by the genome-wide retention of IESs (23), which should
29 disturb the expression of numerous genes. Alternatively, because we noticed that small
30 RNAs, which are normally produced predominantly from IESs, were also produced from
31 non-IES sequences in the *TPB2* mutants (Figure 4B, bottom), we suggest that *Tpb2p* may
32 have a non-catalytic function to confine late-scnRNA production to IESs and the reduction
33 of late-scnRNAs that map Type-B IESs might be due to the gain of late-scnRNA production

1 out of IESs. Nevertheless, the reduction of late-scnRNAs from Type-B IESs in the *TPB2*
2 mutants was approximately two-fold, in contrast to the more than 2,000-fold reduction
3 of IES excision in these mutant cells.

4
5 The results obtained with the *TPB2* mutants do not exclude the possibilities that some
6 late-scnRNA biogenesis occurs after IES excision in wild-type cells and/or late-scnRNA
7 biogenesis normally occurs exclusively after IES excision in wild-type cells but some
8 compensatory mechanism allows the *TPB2* mutants to produce late-scnRNAs without IES
9 excision. However, together with the fact that late-scnRNAs were detected even at the
10 time when elimination of their derived IESs was not detectable in wild-type cells (Figure
11 1), the results shown in this study led us to conclude that the biogenesis of late-scnRNAs
12 occurs prior to the excision of IESs.

13

1 Discussion

2

3 Two technical advances made in this study, the sensitive detection of de novo DNA
4 elimination (Figure 1) and the production of *TPB2*-null mutants (Figure 3), enabled us to
5 show that late-scnRNAs accumulate at least a few hours before the elimination of their
6 derived IESs becomes detectable (Figure 2) and that IES excision is dispensable for their
7 biogenesis (Figure 4). Therefore, secondary small RNAs are produced from IESs prior to
8 their excision in *Tetrahymena*, at least in part. This situation contrast considerably with
9 the excision-dependent secondary small RNA production observed in another ciliate,
10 *Paramecium*, in which concatenated circles formed from excised IESs are suggested to be
11 the templates for the transcription of the precursors of iesRNAs (secondary small RNAs
12 in this ciliate) (28, 35) (Figure 4C). The post-excision production of secondary small RNAs
13 in *Paramecium* is further supported by the fact that injection of DNA fragments into cells
14 during sexual reproduction can induce ectopic DNA elimination at complementary
15 genomic loci (35). In contrast, our attempts to induce ectopic DNA elimination in
16 *Tetrahymena* by introducing linear and circular DNAs have failed. Considering all of these
17 findings together, we conclude that the modes of secondary small RNA biogenesis are
18 fundamentally different between *Paramecium* and *Tetrahymena*. This difference was
19 unexpected because these two ciliates use otherwise very similar molecular mechanisms
20 for DNA elimination.

21

22 The post-excision production of iesRNAs from concatenated IES circles in *Paramecium*
23 might allow “rolling circle transcription” to produce a large amount of small RNAs that is
24 sufficient to target all copies of IESs in the highly polyploid (~800 copies) somatic MAC
25 genome. Such massive amplification of the small RNA signal is probably not necessary in
26 *Tetrahymena*, which exhibits only ~50 copies in the MAC in vegetative stage and 4 to 8
27 copies when IES elimination occurs. Instead, because the IESs of *Tetrahymena* are much
28 longer (median = 2.8 kb) than those of *Paramecium* (median = 51 bp) (6, 42), secondary
29 small RNA production in *Tetrahymena* is necessary to cover IES regions that are not
30 complementary to the primary small RNAs. The fact that heterochromatin formation, but
31 not the downstream IES excision, is necessary for the secondary small RNA biogenesis in
32 *Tetrahymena* (Figure 4B) strongly indicates that the secondary small RNA production
33 mechanism is directly linked to a heterochromatin machinery. Indeed, the cis-spreading

1 of late-scRNA biogenesis via an RNAi-heterochromatin feedback loop plays an
2 important role in IES elimination in *Tetrahymena* (20, 38), which is reminiscent of the
3 heterochromatin-driven production of small RNAs observed in fungi (43), plants (44) and
4 animals (45, 46) and thus likely originated from an ancient TE silencing mechanism. If
5 the above speculations are correct, the remarkable difference in secondary small RNA
6 biogenesis reflects the evolutionary history of genome organization after the two ciliate
7 lineages diverged several million years ago. In this context, it would be interesting to
8 explore the genomes and DNA elimination mechanisms of other ciliates to trace the
9 evolution of TE-targeting small RNA amplification. Such observations would further
10 provide unique views regarding how TEs and their silencing mechanisms drive the
11 evolution of genome structures and vice versa.

1 **Materials and Methods**

2

3 **Strains.** The wild-type *Tetrahymena thermophila* strains B2086 and CU428 were
4 obtained from the *Tetrahymena* Stock Center. The *PDD1* KO strains (W3-3 and W39-1)
5 were described previously (41). The construction of the *TPB2* mutant strains is described
6 in SI Materials and Methods.

7 **DLR-jPCR assay.** Total genomic DNA was extracted using the NucleoSpin Tissue kit
8 (Macherey-Nagel). 5 µg of the genomic DNA was treated with exonuclease V (NEB) for 1
9 h at 37°C, purified using the QIAquick Gel Extraction kit (QIAGEN), and then treated with
10 Plasmid-Safe DNase (Epicentre) for 16 h at 37°C. The remaining circular DNAs were
11 ethanol precipitated and amplified by rolling circle amplification (RCA) using the Illustra
12 TempliPhi Amplification kit (GE Healthcare) according to the manufacturer's
13 instructions, except that the 30°C incubation was performed for more than 18 h. The
14 junctions of circularized IESs were amplified by PCR using the primers listed in SI
15 Appendix Table S1. The positions of the Type-A and Type-B IESs investigated in this study
16 were described previously (20). The positions of *TPB1/6*-dependent IESs were described
17 in Cheng et al. 2016 (26).

18 **Small RNA analyses.** High-throughput sequencing and analyses of small RNAs were
19 performed as described (14), except that radiolabeled 16 nt and 32 nt RNA oligos (5'-
20 GAAUAGUUUAAACUGU-3'; 5'-AUCUUGGUCGUACGCGGAAUAGUUUAAACUGU-3') were
21 used to excise small RNAs and their intermediate products from gels. The raw sequencing
22 data have been deposited under accession no. GSE125297.

23 **Western blotting and Immunofluorescent staining.** Detailed procedures for SDS-
24 PAGE sample preparation, western blotting and immunofluorescent staining are
25 described in SI Materials and Methods .

1 **Acknowledgments.** We acknowledge the MRI facility, member of the national
2 infrastructure France-BioImaging supported by the French National Research Agency
3 (ANR-10-INBS-04), and the NGS unit of Vienna BioCenter Core Facilities. This work was
4 supported by an Advanced Grant from the “Investissements d’avenir” program Labex
5 EpiGenMed (ANR-10-LABX-12-01), an “Accueil de Chercheurs de Haut Niveau [ACHN]”
6 grant (ANR-16-ACHN-0017) from the French National Research Agency to KM.

7
8 **Author Contributions.** MM and KM designed the experiments, MM, TN and KM
9 conducted the experiments, and MM and KM wrote the paper.

10 11 **Figure Legends**

12 **Figure 1. Detection of de novo IES excision events**

13 (A) Linear DNAs in total genomic DNA were digested with exonucleases (digestion of
14 linear DNA [DLD]). Then, the remaining circular DNAs, including IES circles produced by
15 de novo IES excision in the new MAC, were amplified by rolling circle amplification (RCA),
16 and the junction of the DNA circle from an IES locus was detected by PCR (jPCR). The
17 whole procedure is referred to as DLR-jPCR. (B) Wild-type strains were mated, and
18 genomic DNA was extracted at the indicated hour post-mixing (hpm). The circles excised
19 from the indicated IES loci were analyzed by DLR-jPCR with (+) or without (-) RCA. (C)
20 Genomic DNA from 12 hpm was treated as in (B), but the sample was diluted by the
21 indicated factors before jPCR to detect the IES737 locus.

22 23 **Figure 2. Accumulation of late-scnRNAs in wild-type cells**

24 (A–C) Normalized numbers (RPM [reads per million]) of sequenced 26- to 32-nt small
25 RNAs from wild-type cells at indicated time points uniquely mapping to the
26 representative Type-B IES loci are shown as histograms with 50-nt bins. Sky-blue boxes
27 indicate IESs. The fraction of unique sequences that can be mapped by small RNAs in
28 each 50-nt bin is shown at the bottom. (D) Normalized numbers (RPKM [reads per
29 kilobase of unique sequences per million]) of 26- to 32-nt small RNAs from wild-type cells
30 at indicated time points that uniquely matched the 3,293 Type-B IESs are shown as box
31 plots. The horizontal bar in the box indicates the median value. The bars on the top and
32 bottom of the box indicate the minimum and maximum values, respectively, within 1.5x
33 IQR.

34 35 **Figure 3. Production of *TPB2* mutants**

36 (A) A construct expressing Cas9 from the cadmium-inducible *MTT1* promoter and *TPB2*-
37 targeting gRNAs from the constitutive U6 promoter was introduced into the MAC. Some
38 of the Cas9-gRNA complexes are expected to enter the MIC and introduce a mutation at
39 the *TPB2* locus. Immediately next to the *TPB2* locus, there is a chromosome breakage

1 sequence (Cbs), which remains intact in the MIC but is severed to form the ends of two
2 separate chromosomes in the MAC. PCR with primers “a” and “c” specifically amplifies the
3 *TPB2* MIC locus, while PCR with “b” and “c” amplifies both the *TPB2* MAC and MIC loci.
4 (B) Two gRNAs (T1 and T2) were used to target the sequences in the first exon of *TPB2*
5 (marked with yellow). The *TPB2* MIC and MAC loci were amplified, respectively, with the
6 primers “a” and “c” and the primers “b” and “c” shown in (A). Although primers “b” and
7 “c” could amplify both the MAC and MIC loci, polyploid MAC copies were predominantly
8 represented in this PCR assay. The detected insertions and deletions are indicated in red.
9 (C) Total cell proteins were extracted at the indicated time points from conjugation
10 cultures of either wild-type (WT) cells or *TPB2* mutants (*TPB2* mut, cross 1 = #6 x #38,
11 cross 2 = #41 x #44) and analyzed by western blotting using an anti-Tpb2p antibody (top)
12 and an anti- α -tubulin antibody (bottom).
13

14 **Figure 4. Analyses of DNA elimination and late-scrRNA accumulation in *TPB2*** 15 **mutants**

16 (A) Wild-type (WT) and *TPB2* mutant (*TPB2* mut, mating of #41 and #44) strains were
17 mated and their genomic DNA was extracted at 0, 12 and 18 hr post-mixing (hpm). The
18 excised circles from the indicated IES loci were detected by DLR-jPCR. (B) Wild-type (WT),
19 *TPB2* mutant (*TPB2* mut, cross 1 = #6 x #38, cross 2 = #41 x #44) and *PDD1* KO ($\Delta PDD1$)
20 strains were mated and their small RNAs were sequenced. (Top) Normalized numbers
21 (RPKM [reads per kilobase of unique sequences per million]) of 26- to 32-nt small RNAs
22 at 3, 10 and 12 hr post-mixing (hpm) that uniquely matched the 3,293 Type-B IESs are
23 shown as box plots as in Figure 2D. (Bottom) Normalized numbers (RPM [reads per
24 million]) of sequenced 26- to 32-nt small RNAs at 10 hpm uniquely mapping to the
25 indicated Type-B IES loci are shown as histograms with 50-nt bins. Sky-blue boxes
26 indicate IESs. (C) Comparison of the timing of secondary small RNA biogenesis. In
27 *Tetrahymena* (left), primary (1^o) small RNAs fully or partially cover IESs and induce
28 heterochromatin formation, which triggers the production of secondary (2^o) small RNAs.
29 Secondary small RNAs further induce heterochromatin formation and small RNA
30 production. Eventually, heterochromatinized IESs are excised. In *Paramecium* (right),
31 primary small RNAs cover IESs homogeneously. The excised IESs form concatenated
32 circles, which are transcribed to produce secondary small RNAs. Secondary small RNAs
33 further induce IES elimination and small RNA production.

References

1. Czech B, et al. (2018) piRNA-Guided Genome Defense: From Biogenesis to Silencing. *Annu Rev Genet* 52:131–157.
2. Gebert D, Rosenkranz D (2015) RNA-based regulation of transposon expression. *Wiley Interdiscip Rev RNA* 6(6):687–708.
3. Castel SE, Martienssen RA (2013) RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat Rev Genet* 14(2):100–112.

4. Zhou M, Law JA (2015) RNA Pol IV and V in gene silencing: Rebel polymerases evolving away from Pol II's rules. *Curr Opin Plant Biol* 27:154–164.
5. Luteijn MJ, Ketting RF (2013) PIWI-interacting RNAs: from generation to transgenerational epigenetics. *Nat Rev Genet* 14(8):523–534.
6. Hamilton EP, et al. (2016) Structure of the germline genome of *Tetrahymena thermophila* and relationship to the massively rearranged somatic genome. *Elife* 5.
7. Austerberry CF, Allis CD, Yao MC (1984) Specific DNA rearrangements in synchronously developing nuclei of *Tetrahymena*. *Proc Natl Acad Sci USA* 81(23):7383–7387.
8. Saveliev SV, Cox MM (2001) Product analysis illuminates the final steps of IES deletion in *Tetrahymena thermophila*. *EMBO J* 20(12):3251–3261.
9. Saveliev SV, Cox MM (1994) The fate of deleted DNA produced during programmed genomic deletion events in *Tetrahymena thermophila*. *Nucleic Acids Res* 22(25):5695–5701.
10. Chalker DL, Meyer E, Mochizuki K (2013) Epigenetics of ciliates. *Cold Spring Harb Perspect Biol* 5(12):a017764.
11. Noto T, Mochizuki K (2017) Whats, hows and whys of programmed DNA elimination in *Tetrahymena*. *Open Biol* 7(10).
12. Yao M-C, Fuller P, Xi X (2003) Programmed DNA deletion as an RNA-guided system of genome defense. *Science* 300(5625):1581–1584.
13. Chalker DL, Yao MC (2001) Nongenic, bidirectional transcription precedes and may promote developmental DNA deletion in *Tetrahymena thermophila*. *Genes Dev* 15(10):1287–1298.
14. Schoeberl UE, Kurth HM, Noto T, Mochizuki K (2012) Biased transcription and selective degradation of small RNAs shape the pattern of DNA elimination in *Tetrahymena*. *Genes Dev* 26(15):1729–1742.
15. Mochizuki K, Gorovsky MA (2005) A Dicer-like protein in *Tetrahymena* has distinct functions in genome rearrangement, chromosome segregation, and meiotic prophase. *Genes Dev* 19(1):77–89.
16. Malone CD, Anderson AM, Motl JA, Rexer CH, Chalker DL (2005) Germ line transcripts are processed by a Dicer-like protein that is essential for developmentally programmed genome rearrangements of *Tetrahymena thermophila*. *Mol Cell Biol* 25(20):9151–9164.
17. Mochizuki K, Fine NA, Fujisawa T, Gorovsky MA (2002) Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in *tetrahymena*. *Cell* 110(6):689–699.
18. Noto T, et al. (2010) The *Tetrahymena* argonaute-binding protein Giw1p directs a mature argonaute-siRNA complex to the nucleus. *Cell* 140(5):692–703.
19. Noto T, Mochizuki K (2018) Small RNA-Mediated trans-Nuclear and trans-Element Communications in *Tetrahymena* DNA Elimination. *Curr Biol* 28(12):1938–1949.e5.
20. Noto T, et al. (2015) Small-RNA-Mediated Genome-wide trans-Recognition Network in *Tetrahymena* DNA Elimination. *Mol Cell* 59(2):229–242.
21. Liu Y, et al. (2007) RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in *Tetrahymena*. *Genes Dev* 21(12):1530–1545.
22. Taverna SD, Coyne RS, Allis CD (2002) Methylation of histone h3 at lysine 9 targets programmed DNA elimination in *tetrahymena*. *Cell* 110(6):701–711.
23. Cheng C-Y, Vogt A, Mochizuki K, Yao M-C (2010) A domesticated piggyBac transposase plays key roles in heterochromatin dynamics and DNA cleavage during programmed DNA deletion in *Tetrahymena thermophila*. *Mol Biol Cell* 21(10):1753–

- 1762.
24. Vogt A, Mochizuki K (2013) A domesticated PiggyBac transposase interacts with heterochromatin and catalyzes reproducible DNA elimination in *Tetrahymena*. *PLoS Genet* 9(12):e1004032.
 25. Lin I-T, Chao J-L, Yao M-C (2012) An essential role for the DNA breakage-repair protein Ku80 in programmed DNA rearrangements in *Tetrahymena thermophila*. *Mol Biol Cell* 23(11):2213–2225.
 26. Cheng C-Y, et al. (2016) The piggyBac transposon-derived genes TPB1 and TPB6 mediate essential transposon-like excision during the developmental rearrangement of key genes in *Tetrahymena thermophila*. *Genes Dev* 30(24):2724–2736.
 27. Feng L, et al. (2017) A germline-limited piggyBac transposase gene is required for precise excision in *Tetrahymena* genome rearrangement. *Nucleic Acids Res* 45(16):9481–9502.
 28. Sandoval PY, Swart EC, Arambasic M, Nowacki M (2014) Functional diversification of Dicer-like proteins and small RNAs required for genome sculpting. *Dev Cell* 28(2):174–188.
 29. Bouhouche K, Gout J-F, Kapusta A, Bétermier M, Meyer E (2011) Functional specialization of Piwi proteins in *Paramecium tetraurelia* from post-transcriptional gene silencing to genome remodelling. *Nucleic Acids Res* 39(10):4249–4264.
 30. Lepère G, Bétermier M, Meyer E, Duharcourt S (2008) Maternal noncoding transcripts antagonize the targeting of DNA elimination by scanRNAs in *Paramecium tetraurelia*. *Genes Dev* 22(11):1501–1512.
 31. Lhuillier-Akakpo M, et al. (2014) Local effect of enhancer of zeste-like reveals cooperation of epigenetic and cis-acting determinants for zygotic genome rearrangements. *PLoS Genet* 10(9):e1004665.
 32. Baudry C, et al. (2009) PiggyMac, a domesticated piggyBac transposase involved in programmed genome rearrangements in the ciliate *Paramecium tetraurelia*. *Genes Dev* 23(21):2478–2483.
 33. Bischerour J, et al. (2018) Six domesticated PiggyBac transposases together carry out programmed DNA elimination in *Paramecium*. *Elife* 7.
 34. Kapusta A, et al. (2011) Highly precise and developmentally programmed genome assembly in *Paramecium* requires ligase IV-dependent end joining. *PLoS Genet* 7(4):e1002049.
 35. Allen SE, et al. (2017) Circular Concatemers of Ultra-Short DNA Segments Produce Regulatory RNAs. *Cell* 168(6):990–999.e7.
 36. Hoehener C, Hug I, Nowacki M (2018) Dicer-like Enzymes with Sequence Cleavage Preferences. *Cell* 173(1):234–247.e7.
 37. Yao MC, Yao CH (1994) Detection of circular excised DNA deletion elements in *Tetrahymena thermophila* during development. *Nucleic Acids Res* 22(25):5702–5708.
 38. Suhren JH, et al. (2017) Negative Regulators of an RNAi-Heterochromatin Positive Feedback Loop Safeguard Somatic Genome Integrity in *Tetrahymena*. *Cell Rep* 18(10):2494–2507.
 39. Howard-Till R, Tian M, Loidl J (2019) A specialized condensin complex participates in somatic nuclear maturation in *Tetrahymena thermophila*. *Mol Biol Cell*:mbcE18080487.
 40. Shieh AWY, Chalker DL (2013) LIA5 is required for nuclear reorganization and programmed DNA rearrangements occurring during *tetrahymena* macronuclear differentiation. *PLoS One* 8(9):e75337.
 41. Schwoppe RM, Chalker DL (2014) Mutations in Pdd1 reveal distinct requirements for its chromodomain and chromoshadow domain in directing histone methylation and

- heterochromatin elimination. *Eukaryotic Cell* 13(2):190–201.
42. Guérin F, et al. (2017) Flow cytometry sorting of nuclei enables the first global characterization of Paramecium germline DNA and transposable elements. *BMC Genomics* 18(1):327.
 43. Yamanaka S, et al. (2013) RNAi triggered by specialized machinery silences developmental genes and retrotransposons. *Nature* 493(7433):557–560.
 44. Law JA, et al. (2013) Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* 498(7454):385–389.
 45. Klattenhoff C, et al. (2009) The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* 138(6):1137–1149.
 46. Andersen PR, Tirian L, Vunjak M, Brennecke J (2017) A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature* 549(7670):54–59.