
Transgenerational function of *Tetrahymena* Piwi protein Twi8p at distinctive noncoding RNA loci

BRIAN M. FARLEY and KATHLEEN COLLINS

Department of Molecular and Cell Biology, University of California at Berkeley, Berkeley, California 94720-3202, USA

ABSTRACT

Transgenerational transmission of genome-regulatory epigenetic information can determine phenotypes in the progeny of sexual reproduction. Sequence specificity of transgenerational regulation derives from small RNAs assembled into Piwi-protein complexes. Known targets of transgenerational regulation are primarily transposons and transposon-derived sequences. Here, we extend the scope of Piwi-mediated transgenerational regulation to include unique noncoding RNA loci. Ciliates such as *Tetrahymena* have a phenotypically silent germline micronucleus and an expressed somatic macronucleus, which is differentiated anew from a germline genome copy in sexual reproduction. We show that the nuclear-localized *Tetrahymena* Piwi protein Twi8p shuttles from parental to zygotic macronuclei. Genetic elimination of Twi8p has no phenotype for cells in asexual growth. On the other hand, cells lacking Twi8p arrest in sexual reproduction with zygotic nuclei that retain the germline genome structure, without the DNA elimination and fragmentation required to generate a functional macronucleus. Twi8p-bound small RNAs originate from long-noncoding RNAs with a terminal hairpin, which become detectable in the absence of Twi8p. Curiously, the loci that generate Twi8p-bound small RNAs are essential for asexual cell growth, even though Twi8 RNP is essential only in sexual reproduction. Our findings suggest the model that Twi8 RNPs act on silent germline chromosomes to permit their conversion to expressed macronuclear chromosomes. Overall this work reveals that a Piwi protein carrying small RNAs from long-noncoding RNA loci has transgenerational function in establishing zygotic nucleus competence for gene expression.

Keywords: germline development; noncoding RNA; nuclear RNA silencing; piRNA; Piwi

INTRODUCTION

Proteins of the evolutionarily widespread Argonaute (Ago)/Piwi family have broad functions in genome and gene regulation (Meister 2013; Ipsaro and Joshua-Tor 2015). Ago/Piwi proteins derive sequence specificity from tightly bound small RNAs (sRNAs), which are processed from longer transcripts by diverse mechanisms (Ipsaro and Joshua-Tor 2015; Iwasaki et al. 2015). One relatively ubiquitous sRNA biogenesis pathway involves Dicer endonucleolytic processing of double-stranded (ds) RNAs, which can originate from RNA hairpins, annealed transcripts, or dsRNA transcribed by an RNA-dependent RNA polymerase complex (RDRC). Cleavage results in 20–30 nucleotide (nt) sRNAs, which are loaded into Ago/Piwi proteins (Nakanishi 2016). Some Piwi proteins are instead loaded with sRNAs derived from endonucleolytic and exonucleolytic processing of single-stranded transcripts (Kobayashi and Tomari 2016). Some Ago/Piwi proteins have an endonuclease activity that is activated by sRNA loading and base-pairing to a target, resulting in target RNA

cleavage (Kuhn and Joshua-Tor 2013). Ago/Piwi family members that do not retain a nuclease active site instead recruit regulatory factors to the sRNA-paired target. In the cytoplasm, complexes direct post-transcriptional mRNA regulation by translational repression or induced mRNA turnover (Jonas and Izaurralde 2015). In the nucleus, mechanisms of gene repression include cotranscriptional restraint of RNA polymerase II and assembly of transcriptionally repressive chromatin (Castel and Martienssen 2013; Holloch and Moazed 2015).

Piwi proteins expressed in germline and germline-associated cells mediate transgenerational epigenetic repression of transposon mobility (Luteijn and Ketting 2013; Weick and Miska 2014). Despite remarkably different molecular pathways of Piwi-interacting (pi) RNA biogenesis and Piwi RNP mechanisms of action, diverse organisms require Piwi/piRNA-dependent transposon silencing during sexual reproduction (Weick and Miska 2014; Gebert and Rosenkranz

Corresponding author: kcollins@berkeley.edu

Article is online at <http://www.najournal.org/cgi/doi/10.1261/rna.060012.116>.

© 2017 Farley and Collins This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://najournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

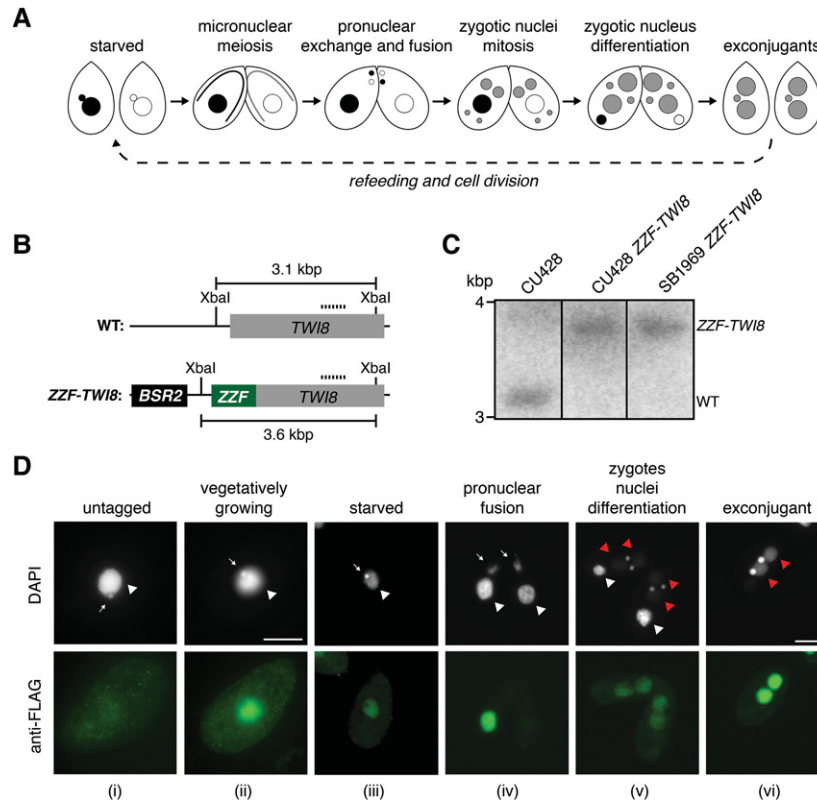


FIGURE 1. Parental *Twis8p* localizes to the zygotic macronucleus. (A) Schematic of *Tetrahymena* conjugation. Starved cells of different mating types (indicated as black or white nuclei) pair and their micronuclei divide by meiosis. One meiotic product from each cell replicates and undergoes a mitotic division and the paired cells exchange one of the resulting pronuclei. After the pronuclei in each cell fuse, each zygotic nucleus replicates and divides twice by mitosis. The two swollen zygotic nuclei differentiate into macronuclei. Parental macronuclei migrate posteriorly and are degraded as zygotic macronucleus differentiation proceeds and paired cells separate as exconjugants. Conjugation is cytologically complete when the exconjugants degrade one of the two micronucleus-sized zygotic nuclei. To enter vegetative growth, exconjugants divide with micronuclear mitosis but no macronuclear replication, generating cells with one micronucleus and one macronucleus. (B) Schematic for *Twis8p* tagging at the *TWI8* locus. In this and all subsequent schematics the restriction nuclease sites indicate the genomic DNA digest and a dashed horizontal line indicates a region detected by a Southern blot probe. *BSR2* indicates the blasticidin resistance cassette; *ZZF* indicates the epitope tag. (C) Southern blot verification of *ZZF-TWI8* strains. *ZZF-TWI8* and WT indicate the genomic DNA fragment that should be detected at targeted and wild-type loci, respectively. All lanes are cropped from the same exposure of the same blot. (D) Immunofluorescence localization of epitope-tagged *Twis8p*. Cells in vegetative growth and starvation were strain *ZZF-TWI8* SB1969. Conjugation was between wild-type CU428 and *ZZF-TWI8* SB1969, which expressed *ZZF-Twis8p* only from the parental macronucleus. DNA was detected by costaining with DAPI, with white arrows marking parental micronuclei, white arrowheads marking parental macronuclei, and red arrowheads marking zygotic macronuclei. The scale bars denote 15 μ m, with the scale bar in (ii) applying to (i) and (ii), and the scale bar in (vi) applying to (iii) through (vi).

2015). Genetic loss of a Piwi protein or piRNA biogenesis enables transposon mobilization, and the ensuing widespread DNA damage leads to sterility. Consistent with these phenotypes, piRNAs include transposon-derived sequences. However, a substantial fraction of piRNAs map to loci that are not of transposon origin or otherwise related to transposon regulation (Le Thomas et al. 2014; Iwasaki et al. 2015). The function(s) of these piRNAs and Piwi RNPs have not been clearly established (Ross et al. 2014; Sarkar et al. 2016).

exconjugants begin vegetative growth, additional DNA amplification generates the 45 C typical of macronuclear chromosomes. The parental macronucleus is degraded prior to the completion of conjugation, and thus each round of sexual reproduction obliges differentiation of a germline nucleus to a somatic nucleus.

Tetrahymena cells express at least six and possibly as many as 12 distinct Piwi family members (Cerutti and Casas-Mollano 2006; Seto et al. 2007; Couvillion et al. 2009). Two

Ciliated protozoa are valuable model organisms for investigating pathways of transgenerational epigenetic regulation (Chalker et al. 2013; Yao et al. 2014; Ruehle et al. 2016). Common to ciliates is the strategy of nuclear dualism: Within a single cell, nuclei are specialized to have either somatic or germline functions. During asexual (vegetative) growth, the diploid germline micronucleus is transcriptionally silent and the polyploid macronucleus is phenotypically expressed. *Tetrahymena* sexual reproduction (conjugation) is induced by starvation and mixing of cells with complementary mating types. Upon cell pairing, micronuclei undergo meiosis, haploid nuclei are exchanged between cells, and gamete fusion creates a zygotic nucleus in each paired cell (Fig. 1A, left). Two mitotic divisions generate four zygotic nuclei in each paired cell, two of which become macronuclei. Paired cells separate as exconjugants (Fig. 1A, right), which if resupplied with nutrients undergo cell division to segregate each new macronucleus to a separate daughter cell. The macronuclear genome contains approximately two-thirds of the sequence complexity of the micronuclear genome, with all transposons and most other repeats removed during macronucleus differentiation in conjugation (Eisen et al. 2006; Coyne et al. 2008). Differentiation of a new macronucleus requires substantial genome remodeling (Yao et al. 2014): chromosome amplification to a copy number (C) of around 12C, sequence-specific chromosome breakage and telomere addition at approximately 180 specific genome locations, and DNA elimination that removes every copy of approximately 6000 different internally eliminated sequences (IES) scattered throughout the micronuclear genome (Yao et al. 1984; Fass et al. 2011). When

Tetrahymena Piwi (Twi) proteins, Twi1p and Twi11p, silence transposable elements in differentiating zygotic macronuclei (Mochizuki et al. 2002; Noto et al. 2015). Early in conjugation, Twi1p is loaded with 28- to 30-nt scan (scn) RNAs generated by the Dicer Dcl1p from annealed nongenic transcripts in the micronucleus (Malone et al. 2005; Mochizuki and Gorovsky 2005). Twi1 RNPs are first imported into parental macronuclei, where scnRNAs complementary to the macronuclear genome are degraded (Schoeberl et al. 2012). Filtered Twi1 RNPs escape the parental macronucleus and relocalize, along with Twi11 RNPs, to developing zygotic macronuclei, where an amplification of IES-targeting scnRNA occurs (Mochizuki et al. 2002; Noto et al. 2015). Ultimately, scnRNA RNPs trigger the trimethylation of histone 3 lysine 9 (H3K9me3) and lysine 27, which recruit chromodomain proteins including Pdd1p that coalesce heterochromatin into punctate foci thought to support IES elimination (Kataoka et al. 2016). H3K9me3 is unique to the developing macronuclei of *Tetrahymena*; remarkably, differentiated macronuclei lack H3K9me3 heterochromatin (Taverna et al. 2002).

Three additional Twi proteins Twi2p, Twi8p, and Twi12p are ubiquitously expressed in vegetative growth as well as conjugation (Couvillion et al. 2009). Twi12p is the most divergent in sequence and in function: it binds mature cytoplasmic tRNAs that are then processed by an unknown endonuclease to leave the CCA-containing tRNA 3' end as the Twi12p-bound sRNA (Couvillion et al. 2010, 2012). Twi12 RNPs are essential for nuclear import of the 5'-3' exonuclease Xrn2p, which mediates ribosomal RNA (rRNA) maturation (Couvillion et al. 2012). On the other hand, Twi2p and Twi8p are loaded more canonically by the *Tetrahymena* Dicer Dcr2p, which generates 23- to 24-nt sRNAs distinct from the conjugation-specific 28- to 30-nt scnRNAs (Lee and Collins 2006, 2007). Dcr2p substrate specificity derives largely from its physical association with three *Tetrahymena* RDRCs, which share the catalytic subunit Rdr1p but have unique combinations of accessory subunits (Rdn1p and Rdf1p, or Rdn1p and Rdf2p, or Rdn2p) that govern cellular template selection (Lee and Collins 2007; Lee et al. 2009; Talsky and Collins 2012). The sRNAs loaded into Twi2p exhibit extreme strand bias, arising only from the RDRC-synthesized strand antisense to cellular transcripts with degenerate or no open reading frame (Lee and Collins 2006; Couvillion et al. 2009). Sequencing of Twi8p-bound sRNAs was less revealing about the specificity of sRNA biogenesis or function, likely due to overexpression of the transgene-encoded tagged Twi8p used for RNP purification (Couvillion et al. 2009). Unlike cytoplasmic Twi2p, Twi8p concentrated in the macronucleus of cells in vegetative growth (Couvillion et al. 2009). A large fraction of Twi8p-bound sRNA was modified by 3'-end 2'-O-methylation (Couvillion et al. 2009), consistent with nuclear localization of the *Tetrahymena* sRNA 2'-O-methylase Hen1p (Kurth and Mochizuki 2009).

To explore nuclear Piwi-protein roles beyond transposon silencing, we investigated *Tetrahymena* Twi8p. *Tetrahymena* transposons are silenced by DNA elimination in differentiating zygotic macronuclei, but Twi8p is constitutively expressed throughout vegetative growth as well as conjugation (Couvillion et al. 2009). We show that Twi8p shuttles from parental macronuclei to the zygotic nuclei committed to macronucleus differentiation. Cells lacking parentally supplied Twi8p arrest at an early stage of somatic macronucleus differentiation, prior to Twi1p- and Twi11p-mediated genome remodeling. Unlike Twi1p- and Twi11p-bound sRNAs, which originate from repetitive elements that are predominantly micronucleus-restricted in sequence, Twi8p-bound sRNAs originate from a small number of loci retained in the macronuclear genome. These loci are essential to cells in vegetative growth, unlike Twi8p. Overall these findings demonstrate a new type of noncoding RNA locus used for sRNA biogenesis and a new role for transgenerational inheritance of Piwi RNPs.

RESULTS

Parentally supplied Twi8p localizes to zygotic macronuclei

A previous study of overexpressed Twi8p indicated that it concentrated in the macronucleus of cells in vegetative growth (Couvillion et al. 2009). To investigate the localization of Twi8p at its endogenous expression level, we tagged the protein encoded at the endogenous gene locus. We targeted the cointegration of a selectable marker (*BSR2* conferring blasticidin resistance) and a Twi8p N-terminal tag with tandem Protein A domains and a triple Flag epitope (the ZZF tag) (Fig. 1B). The targeting construct was introduced by particle bombardment, transformants were selected with blasticidin, and maximal replacement of the endogenous locus was attained by phenotypic assortment (Couvillion and Collins 2012). Because macronuclear chromosomes segregate randomly during cell division, passage of cells into media with serially increasing drug concentration selects for maximal copy number of the recombinant chromosome. Complete replacement of the endogenous *TWI8* locus with the tagged *TWI8* locus was obtained in strains of different mating type, CU428 and SB1969, as determined by genomic DNA Southern blots (Fig. 1C). Fixed cells were stained to detect the Flag tag of ZZF-Twi8p and co-stained with DAPI to detect DNA. Staining background was low in cells lacking ZZF-tagged protein (Fig. 1D, i). Cells expressing ZZF-Twi8p showed predominantly macronuclear localization of the protein in vegetative growth and starvation (Fig. 1D, ii-iii).

To investigate the localization of ZZF-Twi8p across the nuclear dynamics of conjugation (Fig. 1A), we mated wild-type CU428 and ZZF-Twi8p SB1969. Early in conjugation, cell pairs showed staining of only one parental macronucleus (Fig. 1D, iv). At later times, after the creation of zygotic macronuclei, parentally expressed ZZF-Twi8p was detected in

both zygotic macronuclei of both paired cells (Fig. 1D, v), consistent with cytoplasmic mixing across the conjugation junction. As zygotic macronuclei gained ZZF-*Twi8p*, it became undetectable in a parental macronucleus. Because only the parental macronucleus *TWI8* locus expressed tagged protein, not the zygotic macronuclei derived from micronuclei, immunofluorescence localization tracks only parentally expressed *Twi8p*. Therefore, the change in ZZF-*Twi8p* localization strongly suggests that parentally supplied *Twi8p* is transferred to zygotic macronuclei. ZZF-*Twi8p* remained readily detectable in zygotic macronuclei of exconjugants (Fig. 1D, vi), indicating that daughter cells are transgenerationally stocked with parentally supplied *Twi8p* before they begin vegetative growth. The reciprocal cross of wild-type SB1969 and ZZF-*Twi8p* CU428 gave similar results, and the two ZZF-*Twi8p* strains gave viable progeny indicating biological function of the tagged protein (essential function of *Twi8p* in conjugation is described below).

Parentally supplied *Twi8p* is required for completion of conjugation

We next investigated whether parentally supplied *Twi8p* is essential for conjugation. We engineered CU428 and SB1969 strains that replaced the *Twi8p* Piwi-domain open reading frame with a blasticidin resistance cassette (Fig. 2A). Following selection in increasing concentrations of blasticidin to drive maximal assortment of the recombinant chromosome, individual cells were isolated and expanded by culture in the absence of blasticidin to allow backassortment of any remaining wild-type chromosome. Assortment of the modified allele was determined by Southern blot of genomic DNA. Complete replacement of the *TWI8* locus was observed in both strain backgrounds (Fig. 2B), consistent with the nonessential function of *Twi8p* in vegetative growth (Couvillion et al. 2009).

We starved and mated Δ *TWI8* cells to each other or to wild-type cells, in comparison to the control cross of wild-type partners. Cells formed pairs efficiently with or without parental *Twi8p*: >85% of cells were paired by 4 h post-mixing. The progression of conjugation (Fig. 1A) was followed by DAPI staining of parental and zygotic nuclei. Five categories of nuclear arrangement were scored (Fig. 2C shows representative images): paired cells with micronuclear meiosis, paired cells with multiple zygotic nuclei, paired cells with two of the zygotic nuclei committed to become new macronuclei (red arrowheads), exconjugants that contain two macronuclei and two micronuclei, and exconjugants with two macronuclei and one micronucleus that have completed the nuclear program of conjugation.

By 4 h post-mixing, paired cells from all crosses had progressed to micronuclear meiosis (Fig. 2D; gray bars are the cross of two wild-type strains, orange bars are the cross of Δ *TWI8* to wild type, and red bars are the cross of two Δ *TWI8* strains). Also, paired cells from all crosses underwent

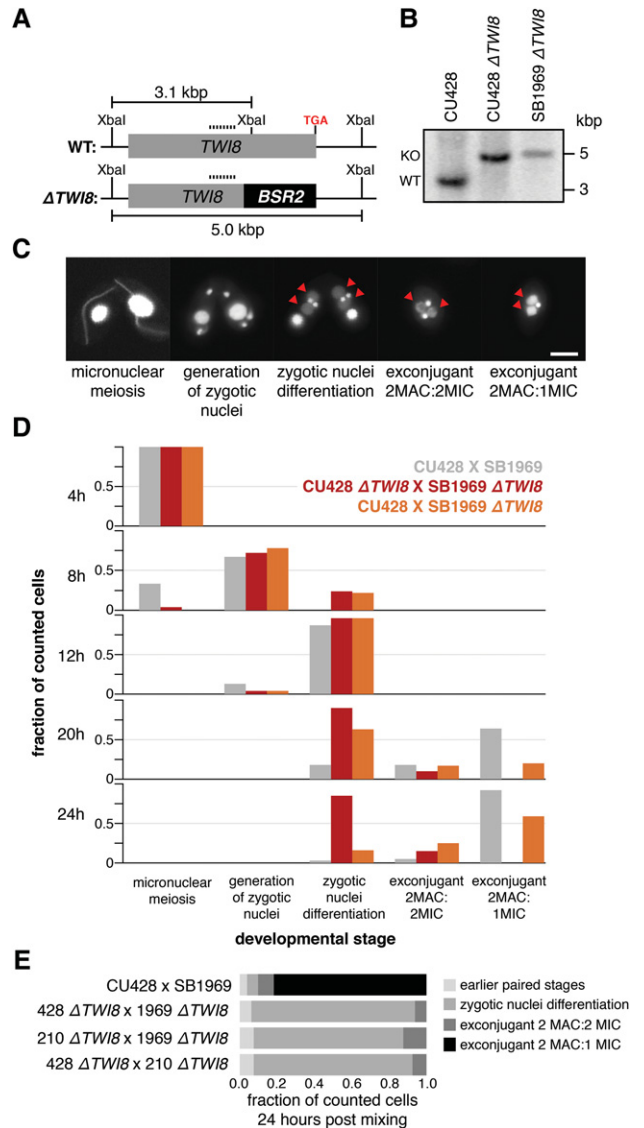


FIGURE 2. Cells lacking parentally supplied *Twi8p* arrest in sexual reproduction. (A) Schematic of *TWI8* disruption strategy. The stop codon and approximately half of the coding sequence of *TWI8* were replaced with a drug resistance cassette (*BSR2*). The dashed horizontal line indicates the region used to synthesize radiolabeled probe for the Southern blot in panel B. (B) Southern blot verification of Δ *TWI8* strains. WT and KO indicate the positions of digested genomic DNA expected for wild-type and targeted loci, respectively. (C) Representative images of DAPI-stained cells. Each of the stages used to score conjugation progression in panel D is shown, arranged in chronological order from left to right. The scale bar denotes 15 μ m. (D) Progression of nuclear configurations was monitored by time-course (h indicates hours post-mixing). Cells were assigned to categories on the basis of nuclei configurations, as described in the text. Each bar graph represents fraction of cells in each category at each time point. The 20-h time point was collected from parallel crosses performed with the same cultures of starved cells. At least 100 cells were counted from each cross at each time point. (E) Progression of nuclear configurations was monitored at 24 h post-mixing for *TWI8* knockouts in various wild-type strain backgrounds. Cells were mixed, collected, fixed, stained, and counted as in panel D. At least 100 cells were counted from each cross. Some labeling in this and subsequent figures uses 428 to indicate CU428 and 1969 to indicate SB1969.

two rounds of zygotic nucleus mitosis to generate two small, DAPI-bright nuclei that retain micronuclear identity and two swollen, DAPI-poor nuclei committed to new macronucleus identity (Fig. 2C, center; all four zygotic nuclei are clustered anterior to the parental macronucleus, and nuclei committed to macronuclear differentiation are indicated with red arrowheads). Lack of parental *Twi8p* did not alter progression to this stage of conjugation (Fig. 2D, see scoring at 12 h). Pairs of wild-type cells or paired wild-type and $\Delta TWI8$ cells continued to the end-point of conjugation, which in nutrient-poor starvation medium is the generation of exconjugants with one micronucleus and two macronuclei (Fig. 2C and 2D gray and orange bars). The parental macronucleus was resorbed completely, with all transcription dependent on the new zygotic macronuclei. In contrast, paired $\Delta TWI8$ cells arrested mid-way through conjugation (Fig. 2D, red bars). The $\Delta TWI8$ pairs retained two small, micronucleus-progenitor zygotic nuclei as well as two swollen, macronucleus-progenitor zygotic nuclei, and the vast majority retained parental macronuclei (Fig. 2C and 2D, stages of zygotic nuclei differentiation and 2MAC:2MIC). To confirm that the $\Delta TWI8$ phenotype is not specific to one strain background, we performed crosses with cells from a $\Delta TWI8$ SB210 background as well. Twenty-four hours after mixing of nutritionally starved cells, pairs from mating of SB210 $\Delta TWI8$ with CU428 $\Delta TWI8$ or SB1969 $\Delta TWI8$ had a conjugation arrest parallel to that observed for mating of CU428 $\Delta TWI8$ and SB1969 $\Delta TWI8$ (Fig. 2E).

Because the $\Delta TWI8$ conjugation arrest was rescued by mating to a wild-type partner (Fig. 2D, orange bars; reciprocal cross results were similar, data not shown), it is unlikely that *Twi8p* is required for parental micronuclear genome stability in vegetative growth. This conclusion was reinforced by confirming the functionality of zygotic macronuclei in the progeny of $\Delta TWI8$ and wild-type cell conjugation by successful growth in cycloheximide, which is dependent on macronuclear expression of a gene present only in the SB1969 micronucleus (see Materials and Methods). However, the $\Delta TWI8 \times$ wild-type cross did generate more exconjugants with 2MAC:2MIC content instead of 2MAC:1MIC (Fig. 2D, 24 h). Therefore, some $\Delta TWI8 \times$ wild-type cell pairs may not have enough diffusion across the cytoplasmic conjugation junction to fully support *Twi8p* function.

Lack of parental *Twi8p* arrests macronucleus differentiation prior to genome remodeling

To address the cellular function of *Twi8p*, we next investigated what known events of zygotic macronuclear genome differentiation do or do not occur in the absence of parental *Twi8p* (Fig. 3A). Within 4 h post-mixing, transcription in parental micronuclei produces dsRNA that is processed to *Twi1p*-bound 28- to 30-nt scnRNA (Mochizuki et al. 2002). Because *Twi8p* remains localized in parental macronuclei during this early interval of conjugation (Fig. 1D, ii), *Twi8p*

function would not be expected to influence scnRNA biogenesis. Indeed, both wild-type and $\Delta TWI8$ paired cells generated abundant 28- to 30-nt scnRNAs (data not shown). The scnRNA-directed formation of heterochromatin begins in paired cells with a nuclear configuration similar to that of the $\Delta TWI8$ conjugation arrest. We therefore tested for the presence of H3K9me3 in developing zygotic macronuclei by immunofluorescence staining of paired cells at 10 h post-mixing. Paired cells from all wild-type and $\Delta TWI8$ strain combinations had zygotic macronuclei with comparable detection of H3K9me3 (Fig. 3B) and H3K27me3 (data not shown). Thus, the developing zygotic macronuclei of $\Delta TWI8$ paired cells are competent to establish these histone modifications.

We also examined zygotic macronucleus import and distribution of *Pdd1p*, a chromodomain protein that binds H3K9me3 and induces DNA elimination (Madireddi et al. 1996; Taverna et al. 2002; Liu et al. 2007). Paired cells from all wild-type and $\Delta TWI8$ strain combinations at 14 h post-mixing had imported *Pdd1p* to developing zygotic macronuclei (Fig. 3C). However, only cell pairs with a wild-type partner had the punctate foci of *Pdd1p* coincident with excision of chromatin targeted for degradation (Fig. 3C, top and bottom panels). Foci of *Pdd1p* were not detected in the zygotic macronuclei of $\Delta TWI8$ cells (Fig. 3C, middle panels), which at 14 h post-mixing remained paired while pairs containing a wild-type partner separated to exconjugants. We conclude that the conjugation arrest of $\Delta TWI8$ cells occurs after zygotic macronuclei are functional for protein import but before the assembly of heterochromatin bodies required for DNA elimination.

To complement the cytological assays described above, we used a molecular assay to examine whether the $\Delta TWI8$ conjugation arrest occurred prior to DNA elimination. The micronuclear M-element locus has alternative IES elimination boundaries (Fig. 3D, right): Maximal DNA elimination removes ~900 bp ($\Delta M-0.9$), while an alternative excision removes ~600 bp ($\Delta M-0.6$) (Austerberry et al. 1984). Using PCR primers that anneal to sequences flanking the boundaries of maximal IES elimination, we amplified M-element loci in the genomic DNA of parental strains or conjugated cells at 30 h post-mixing. Strains CU428 and SB1969 both have parental macronuclear M-element loci of only the $\Delta M-0.6$ allele (Fig. 3D, lanes 1–3; note that micronuclear loci are at a much lower copy number and the long PCR product from an intact M-element would be disfavored in amplification). After conjugation of wild-type CU428 and SB1969, progeny macronuclei had both $\Delta M-0.6$ and $\Delta M-0.9$ alleles (Fig. 3D, lane 4), indicating successful IES elimination in zygotic macronuclei. Conjugation of wild-type and $\Delta TWI8$ cells also generated the $\Delta M-0.9$ allele (Fig. 3D, lanes 6,7), indicative of zygotic genome IES elimination. In contrast, conjugation of two $\Delta TWI8$ cells did not (Fig. 3D, lane 5).

Developing macronuclei undergo chromosome breakage in addition to DNA elimination. To test whether $\Delta TWI8$ conjugation arrest occurred prior to chromosome breakage, we assayed for a transient post-breakage form of the chromosome

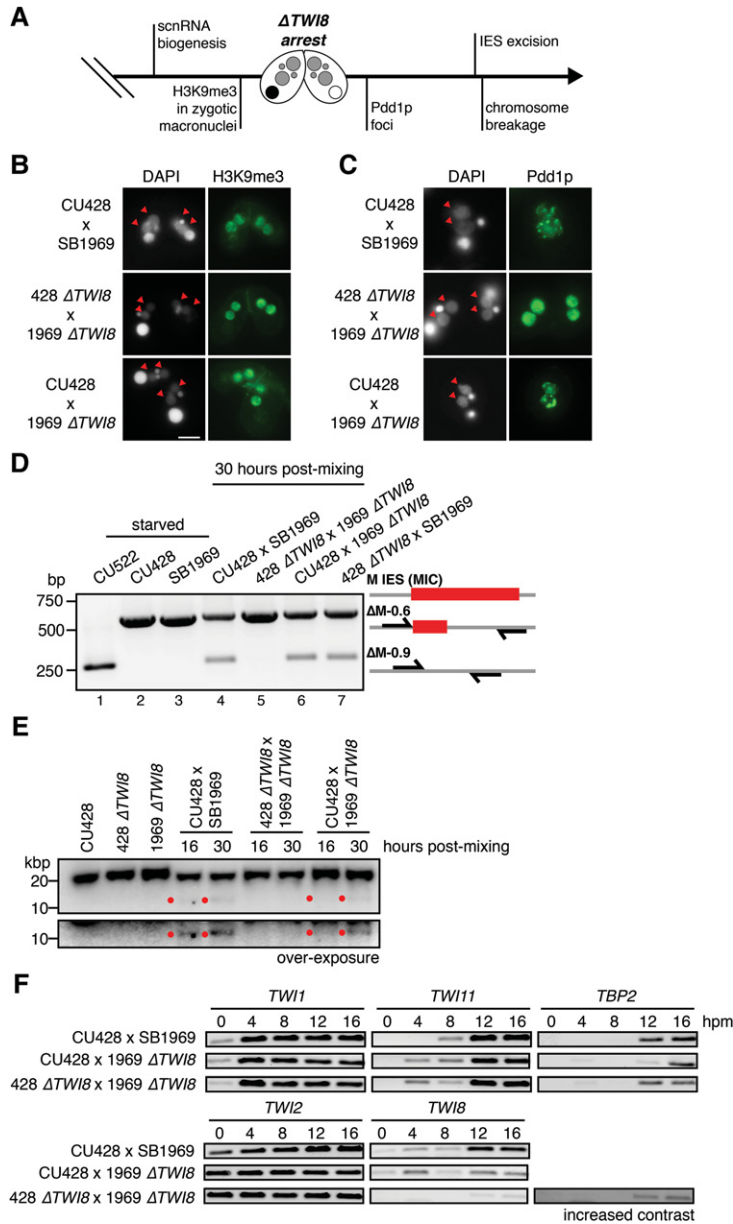


FIGURE 3. Cells lacking parentally supplied *Twi8p* do not undergo zygotic macronucleus differentiation. (A) Approximate timeline for events during conjugation. (B) Immunofluorescence localization of H3K9me3. Cells were collected in conjugation at 10 h post-mixing, fixed, and stained for H3K9me3 with a DAPI counterstain. Red arrowheads in the DAPI-stained panels indicate zygotic macronuclei. The scale bar denotes 15 μ m. (C) Immunofluorescence localization of Pdd1p. Cells were collected in conjugation at 14 h post-mixing, fixed, and stained for Pdd1p with a DAPI counterstain. Red arrowheads in the DAPI-stained panel indicate zygotic macronuclei. The scale bar denotes 15 μ m. (D) PCR-based assay for M-element IES excision. The macronuclear configurations of the M-element locus were assayed by PCR of total genomic DNA. Δ M-0.6 and Δ M-0.9 refer to the alleles produced after elimination of \sim 600 or \sim 900 bp of the micronuclear M-element locus. The schematic at right indicates the PCR primers used (black half-arrows) and the maximal IES (red). (E) Southern blot detection of chromosome breakage at the micronuclear rDNA locus. Undigested genomic DNA was resolved on an agarose gel and transferred to a nylon membrane for Southern blotting using an rDNA-specific probe. The bottom panel is an increased contrast of the 11-kbp region of the panel above. Red dots mark the position of the 11-kbp intermediate of rDNA chromosome formation generated by chromosome breakage. (F) RT-PCR detection of mRNA expression. Total RNA was treated with DNase prior to cDNA synthesis initiated by an oligodeoxythymidine primer. Each transcript was amplified with a gene-specific primer pair. *TWI2* is expressed throughout the *Tetrahymena* lifecycle and was used as a cDNA quality and recovery control. The increased contrast panel is of *TWI8* mRNA in Δ *TWI8* conjugating cell pairs.

that encodes the rRNA precursor transcribed by RNA Polymerase I (designated the rDNA chromosome). The micronuclear genome contains a single copy of the locus for RNA polymerase I transcription of rRNA. During conjugation, this 11-kbp region is excised by chromosome breakage, followed by conversion to the palindromic 21-kbp rDNA chromosome of cells in vegetative growth (Yao et al. 1987). Because the 11-kbp intermediate of rDNA chromosome formation is unique to conjugating cells, it is a discriminating indicator of zygotic macronuclear chromosome breakage (Cranert et al. 2014). We collected genomic DNA from conjugation of wild-type cells, wild-type with Δ *TWI8*, or Δ *TWI8* cells at 0, 16, and 30 h post-mixing and assayed for the rDNA processing intermediate by Southern blot (Fig. 3E). Paralleling the findings for M-element IES elimination described above, the 11-kbp rDNA product of chromosome breakage was detected upon conjugation of wild-type cells or wild-type with Δ *TWI8* but not Δ *TWI8* cells (Fig. 3E). Overall, we conclude that lack of parental *Twi8p* arrests new macronucleus differentiation prior to both chromosome breakage and IES elimination.

Lack of parental *Twi8p* does not preclude transcription from zygotic macronuclei

Developing macronuclei support mRNA transcription before genome rearrangement (Wenkert and Allis 1984). For example, zygotic nuclei express the *TW11* and *TBP2* mRNAs that encode proteins required for DNA elimination and chromosome breakage (Cheng et al. 2010; Noto et al. 2015). In comparison, *TW11* mRNA is expressed both from parental macronuclei in the first few hours of conjugation and in zygotic macronuclei at about the onset of IES elimination (Mochizuki et al. 2002; Noto et al. 2015). To investigate whether lack of parental *Twi8p* arrested zygotic macronucleus differentiation prior to the onset of mRNA transcription, we collected total RNA from conjugations of wild-type cells, wild-type with Δ *TWI8*, and

$\Delta TWI8$ cells at 0, 4, 8, 12, and 16 h post-mixing. Total RNA was reverse-transcribed with poly-deoxythymidine primer to generate cDNA for detection of individual mRNAs by RT-PCR with specific primer pairs.

Twii expression from parental macronuclei occurred in all cell pairings, evident in the comparable level of *Twii* RT-PCR product detected at 4 h post-mixing (Fig. 3F). *Twiii* and *TBP2* mRNA levels were also comparable in all pair combinations, increasing at 12 h post-mixing (Fig. 3F), as expected for transcription from zygotic macronuclei. We used the ubiquitously expressed *Twii2* mRNA as a cDNA quality control (Fig. 3F). Some *Twii8* mRNA expression was detected in paired $\Delta TWI8$ cells at 12 and 16 h post-mixing (Fig. 3F), which is necessarily from the zygotic macronuclei that have wild-type *Twii8* versus the parental macronuclei that are $\Delta TWI8$. However, *Twii8* mRNA level was much lower without expression from parental as well as zygotic developing macronuclei, and even pairs with one wild-type and one $\Delta TWI8$ partner had less *Twii8* mRNA accumulation than pairs of cells with *Twii8* mRNA expressed from both parental macronuclei (Fig. 3F). That *Twii8* mRNA transcription is predominantly from parental rather than zygotic nuclei is consistent with the failure of zygotic *Twii8* mRNA transcription to relieve the conjugation arrest of parental macronucleus $\Delta TWI8$ cell pairs. In combination, all of the assays described above suggest that $\Delta TWI8$ conjugation arrest occurs after zygotic nuclei gain transcriptional competence but before the conversion of germline to somatic genome structure.

Twii8p-bound sRNAs derive from specific noncoding RNA loci

Piwi protein function is guided by bound sRNAs. To characterize Twii8p-bound sRNAs, we purified ZZF-Twii8p from the SB1969 and CU428 strain backgrounds in vegetative growth or starvation, or from conjugations of ZZF-Twii8p SB1969 and wild-type CU428 or ZZF-Twii8p CU428 and wild-type SB1969 sampled at 24 h post-mixing. Twii8 RNPs were purified from cell extracts using a first step of binding to IgG agarose and elution with TEV protease and a second step of binding to Flag antibody resin and elution with Flag peptide. A single protein was prominent by silver staining, which was specific to purifications from ZZF-Twii8p cell extracts versus cell extracts lacking tagged protein (Fig. 4A shows a representative purification). Co-enriched RNAs were extracted, resolved by denaturing polyacrylamide gel electrophoresis (PAGE), and visualized by SYBR Gold staining. Consistent with overexpressed Twii8p purification (Couvillion et al. 2009), endogenous-level Twii8p was bound only to 23- to 24-nt sRNAs (Fig. 4B shows a representative purification).

Gel-purified sRNAs were cloned for high-throughput sequencing, and strand-specific sRNA reads were mapped to the macronuclear and micronuclear genomes (see Materials and Methods). Unlike the 28- to 30-nt scnRNAs enriched in micronucleus-restricted IES sequences, Twii8p 23- to 24-

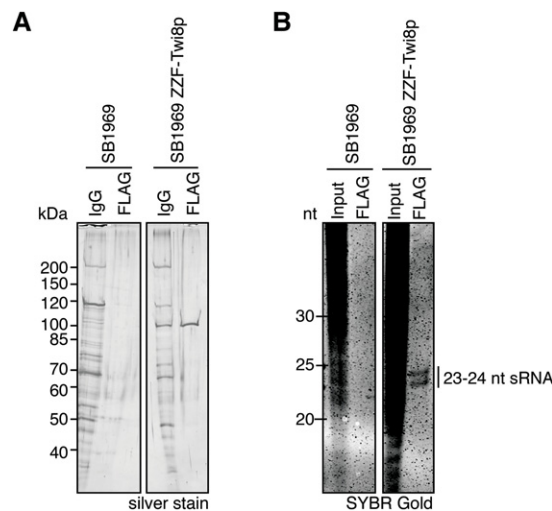


FIGURE 4. Twii8p copurifies 23- to 24-nt sRNAs. (A) Representative purification of Twii8p detected by silver staining after SDS-PAGE. Purifications were performed in parallel from extracts of cells in vegetative growth. Protein profile is shown after the first step of purification (eluted from IgG agarose) or the second step (eluted from Flag antibody resin). (B) SYBR Gold visualization of sRNAs isolated from the purifications in A. RNA samples are from a small amount of starting cell extract (Input) or after the completion of second-step purification (Flag).

nt sRNAs from cells in growth, starvation, and conjugation mapped to the macronuclear genome. To assign genomic loci from which sRNAs originated, we grouped reads that mapped within 500 bp of another read as a single locus and normalized the total number of reads mapped to each locus by the predicted length of the locus (reads per kbp of locus per million reads, or RPKM). We combined reads from vegetative growth or starvation or conjugation and compared bound sRNAs across the *Tetrahymena* life cycle (Fig. 5A). The population of Twii8p-bound sRNAs remained remarkably consistent in sequence composition despite dramatic changes in gene expression and scnRNA production.

Seven macronuclear loci accounted for over 95% of mapped reads from each sRNA library (Fig. 5A, dark-fill circles). We therefore describe these loci as Twii8p-associated sRNA loci (TASLs). Reads mapped to both DNA strands (Fig. 5B; red indicates sense-strand and blue indicates anti-sense-strand, see below). Gaps in read mapping are consistent with sRNA biogenesis from dsRNA synthesized after splicing of a primary transcript, with the TASL-4 locus harboring separate TASL-4A and TASL-4B transcripts (Fig. 5B). Only TASL-6 contains a possible open reading frame longer than 30 amino acids. The TASLs also lack nearby transcribed genes, which is highly atypical for the *Tetrahymena* macronucleus (Coyne et al. 2008). Only TASL-7 is within 500 bp of a possible open reading frame (Fig. 5B, gray bar), but that open reading frame lacks any annotated evidence or suggestion of function. We conclude that TASL loci generate noncoding RNAs from otherwise transcriptionally isolated regions of the macronuclear genome.

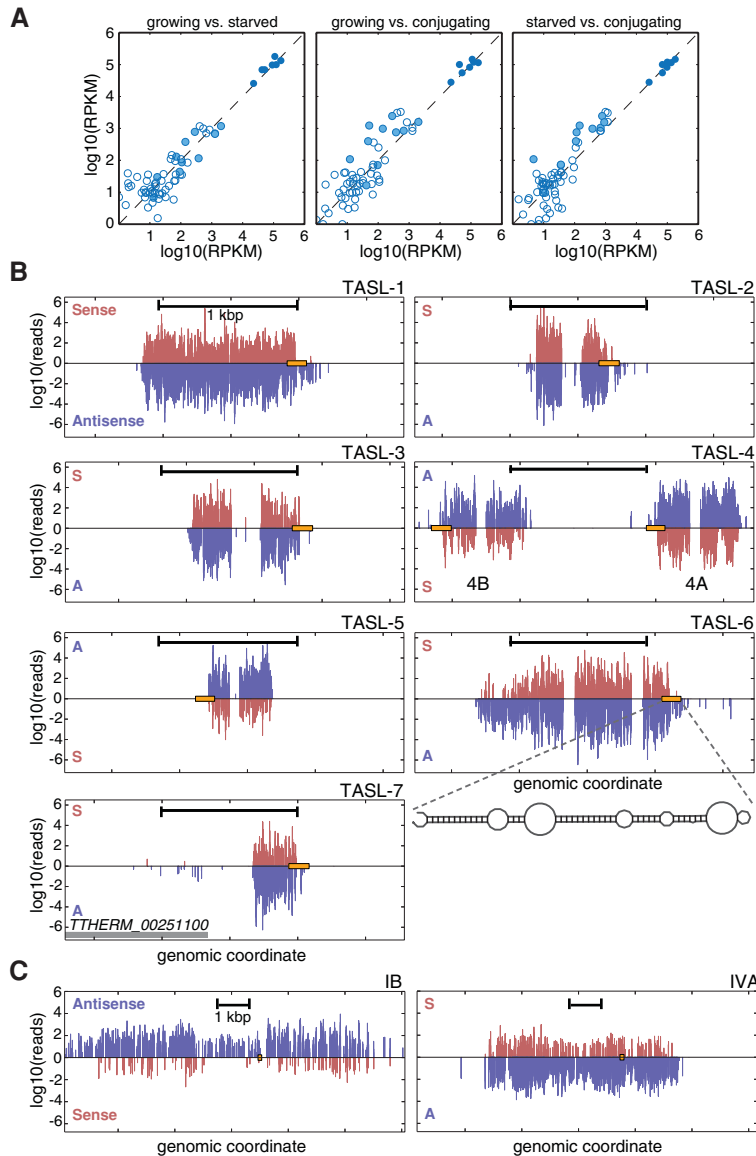


FIGURE 5. *Twi*8p-bound sRNA reads map to noncoding loci in the macronuclear genome. (A) Correlation plots of normalized read counts for the 50 loci with maximal read mapping. Labels above the plots indicate the growth conditions plotted on the x- and y-axes, respectively. Dark blue circles indicate TASLs and light blue circles indicate PGC loci. The dashed line indicates plot positions expected for loci with identical sRNA levels between conditions. (B) Distribution of mapped reads at the seven loci producing the highest normalized abundance of *Twi*8p-bound sRNAs. The coordinate of each read 5' end is plotted as a function of the genomic coordinate for each TASL. Reads mapping to the plus strand of the macronuclear genome are plotted as positive values, while reads mapping to the minus strand are plotted as negative values. Reads antisense to predicted TASL primary transcripts are plotted in blue, while reads from the predicted sense-strand transcripts are plotted in red. Read counts are the sum of six data sets. Black bars represent a scale length of 1 kbp. A predicted open reading frame is indicated by gray rectangle labeled with the annotated gene name. Orange rectangles denote the location of predicted hairpins in the sense-strand transcripts. The predicted secondary structure of the representative TASL-6 RNA hairpin is illustrated. (C) Distribution of mapped reads at PGC loci of two pseudogene families (I and IV), which are representative of other PGC, depicted as described in B.

Five of the seven TASLs (excepting TASL-2 and TASL-6) overlap previously identified “phased cluster loci” that generate abundant sRNAs associated with *Twi*2p (Couvillion et al.

2009). *Twi*2p-bound sRNAs are antisense to a primary transcript detectable by RT-PCR (Couvillion et al. 2009). Unlike *Twi*2-bound sRNAs, *Twi*8p-bound sRNAs from the TASLs are produced from both strands and without 5'-end phasing (Fig. 5B, Supplemental Fig. S1). At every TASL, the *Twi*8p sense-strand sRNA reads tapered off near or within a predicted primary-transcript hairpin of >30 bp (Fig. 5B, shown for TASL-6). This hairpin is proposed to be a determinant of RDRC recruitment, either directly or indirectly, by inducing endonucleolytic cleavage to generate a transcript 3' end for the loop-back initiation by a *Tetrahymena* RDRC (Talsky and Collins 2010).

Beyond the TASLs, among the 50 loci with maximal read coverage density of *Twi*8p-bound sRNA, 10 of the loci are previously identified pseudogene clusters (Fig. 5A, light-fill circles). Pseudogene clusters (PGC) generate abundant cellular 23- to 24-nt sRNAs with the strand asymmetry hallmark of loading on *Twi*2p (Lee and Collins 2006; Couvillion et al. 2009). Suggestive of pseudogene origin, PGC transcripts share a degenerate open reading frame and genome-encoded polyadenosine tract on the sense strand of the primary transcript detectable by RT-PCR (Lee and Collins 2006). Unlike *Twi*2p-bound PGC sRNAs, *Twi*8p-bound PGC sRNAs are produced from both strands (Fig. 5C). However, *Twi*8p-bound PGC sRNAs and *Twi*8p-bound sRNAs from some TASLs do have an overall antisense bias that is quantitatively variable across individual loci (Table 1). This asymmetry reflects a combination of *Twi*8p loading by the *Twi*8p-specific pathway without strand specificity and some cross-loading by a strand-selective *Twi*2p loading pathway (see Discussion). A third pathway may load a small fraction of *Twi*8p by Dcr2p cleavage of annealed rather than RDRC-synthesized dsRNA, as detected previously (Couvillion et al. 2009) and here (Table 1, *Twi*8.1).

***Twi*8p reduces TASL transcript levels**

Because nuclear Piwi proteins reduce transcription of their target loci, we investigated the change in accumulation of

TABLE 1. Strand bias of Twi8p sRNAs from the most abundant sRNA-producing loci

Locus	Antisense bias	Scaffold	Start	End
TASL-1	6.6	scf_8254594	879806	881242
TASL-2	3.4	scf_8254370	141562	142534
TASL-3	8.3	scf_8254594	261231	262485
TASL-4	6.7	scf_8254597	3756	6190
TASL-5	37	scf_8254638	467098	467708
TASL-6	19	scf_8254470	647878	650081
TASL-7	50	scf_8254590	192790	194281
IA	29	scf_8254028	22716	28008
IB	46	scf_8254600	158161	170065
IC	48	scf_8254638	960882	969246
ID	93	scf_8254659	1291175	1294032
IIA	79	scf_8254557	43710	48085
IIB	75	scf_8254822	133026	138375
IVA	23	scf_8254572	2710	9902
IVB	46	scf_8254718	114857	123998
IVC	62	scf_8254631	1528	8420
IVD	53	scf_8254233	1682	4307
Rpt1	36	scf_8254678	16064	20818
Twi8.1	150	scf_8254552	88977	89131
Twi8.4	1	scf_8254428	46446	53361
Twi8.6	5.1	scf_8254486	383033	384878

Locus is the name of the locus as defined here (TASLs) or in Couvillion et al. (2009) (others). Antisense bias is the ratio of sRNA counts derived from the antisense strand to sRNA counts from the sense strand observed in the data described in Figure 5. Scaffold, start, and end describe the genomic location of the associated locus as inferred from the sequencing of Twi8p-associated sRNAs.

TASL transcripts in the absence of Twi8p. We prepared total RNA from conjugation of two wild-type or two $\Delta TWI8$ strains at 0, 8, or 12 h post-mixing. Wild-type starved cells or conjugating cell pairs accumulated only TASL-6 transcripts barely detectable by Northern blot (Fig. 6, lanes 7–9). In contrast, in $\Delta TWI8$ cells, transcripts were detected from all TASLs except TASL-7 (Fig. 6, lanes 4–6 and 10–12). In starved $\Delta TWI8$ cells, four TASL transcripts had increased abundance (Fig. 6, lane 4; TASLs 1, 3, 4, and 5). These decreased in abundance at 12 h post-mixing (Fig. 6, lane 6). On the other hand, transcripts from TASL-2 and TASL-6 were maximal in $\Delta TWI8$ cells at 12 h post-mixing (Fig. 6, lane 12). A relatively low level of TASL-2 and TASL-6 transcription in growing cells would account for why sRNAs from these two TASLs were less abundant in the total cellular and Twi2p-bound sRNA pools (Couvillion et al. 2009). Overall, we conclude that Twi8p limits the expression or accumulation of TASL long-noncoding (nc) RNA primary transcripts.

Macronuclear TASLs are required for vegetative growth

Using methods similar to those described for production of $\Delta TWI8$ cells above, individual TASLs were targeted for deletion in the macronuclear genome of strains CU428 and SB1969. We attempted to individually replace representative TASLs (TASL-2, TASL-3, and TASL-6) with a neomycin re-

sistance cassette, removing the entire region of mapped sRNA reads and ~ 300 bp on each side (Fig. 7A). Successful TASL targeting was detected by Southern blot of genomic DNA, but the targeted chromosome failed to replace the wild-type chromosome even under high selective pressure (Fig. 7B). In our experience, an essential protein-coding gene will typically tolerate loss of copy number to about half of wild-type, much more than was attained for any TASL knockout over repeated attempts including use of an alternate selection cassette (data not shown). We conclude that the macronuclear TASLs are essential for vegetative growth.

Curiously, the TASLs are distributed across all five micro-nuclear chromosomes. To test the hypothesis of TASL ncRNA transcript function in the micronucleus, we attempted to knock out individual TASLs in cells without a micronucleus that lack the micronucleus integrity checkpoints. Complete lack of micronuclear chromosomes in the amicro-nucleate strain BI3840 allowed us to test macronuclear chromosome assortment by multiplex PCR. One primer pair in each reaction amplified *TBPI* as a normalization control for genomic DNA input (Fig. 7C, control). The second primer pair amplified a portion of the targeted TASL that would be replaced by the drug resistance cassette. Therefore, the amount of wild-type chromosome retention would be reflected in the amount of PCR product. To establish the sensitivity of the assay, we performed PCR with genomic DNA from TASL-2 ΔUP cells, using primers within the upstream region that was eliminated from the macronucleus (see below). PCR signal for the TASL decreased substantially because only the wild-type TASL in the micronucleus provided template for PCR amplification (Fig. 7C). For the TASL knockout attempts in the amicro-nucleate strain, none were successful in substantially replacing the TASL with a drug resistance cassette (Fig. 7C). These findings suggest that the essential function of TASLs in vegetative growth does not involve micronuclear chromosomes.

Next, we attempted to eliminate specific regions within a TASL by gene replacement (Fig. 7D). We generated targeting constructs that replaced either 800 bp upstream of the

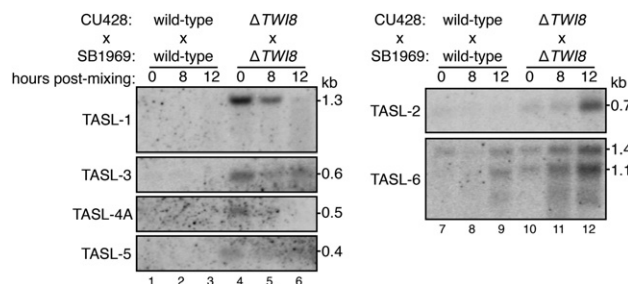


FIGURE 6. TASL long-ncRNAs have increased accumulation in the absence of Twi8p. Northern blots of TASL transcript accumulation across time courses of wild-type or $\Delta TWI8$ cell conjugation. The sizes of transcripts detected by Northern blot roughly match the length of the read mapping regions.

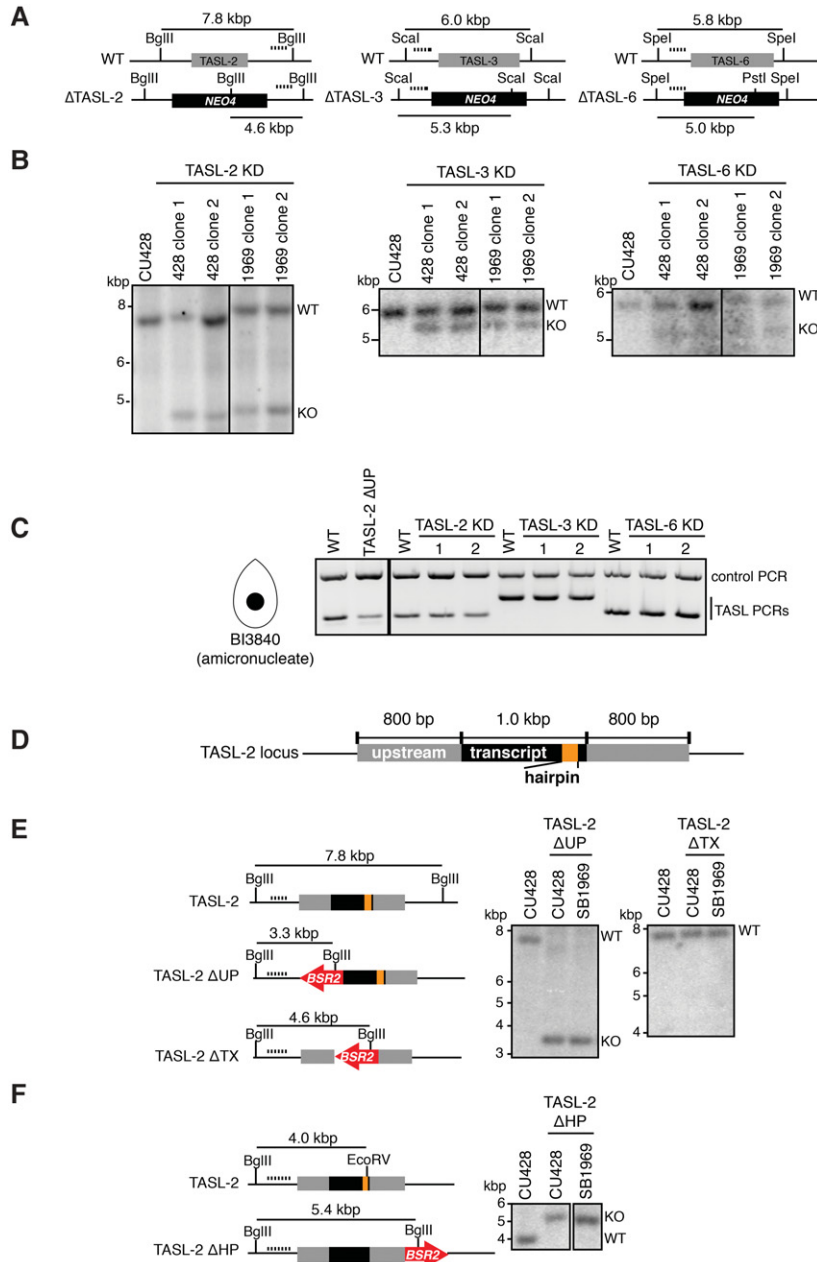


FIGURE 7. TASLs are genetically essential for vegetative cell growth. (A) Schematics of TASL disruptions. *NEO4* indicates a paromomycin resistance cassette. (B) Southern blots detecting failure of macronuclear TASL knockout. KD indicates knockdown rather than knockout of each macronuclear TASL in independently selected cells isolated as clonal cell lines prior to release from selection. WT and KO indicate the positions of digested genomic DNA expected for wild-type and targeted loci, respectively. (C) Failure of macronuclear TASL knockout in a strain lacking a micronucleus. Following drug selection and back-assortment, genomic DNA was used as a template for multiplex PCR reactions containing primer pairs for a control locus (*top* product) and the targeted TASL (*bottom* product). Loss of TASL copy number would reduce the TASL PCR product. (D) Schematic of *TASL-2* regions targeted for deletion. (E,F) Schematics and Southern blots of macronuclear *TASL-2* targeting for upstream, transcript, or transcript hairpin deletion.

predicted *TASL-2* transcript (*TASL-2* Δ UP) or the predicted transcript itself (*TASL-2* Δ TX) with a blasticidin resistance cassette (Fig. 7E, left). We also generated a targeting construct that removed the predicted hairpin of the *TASL-2* transcript

(*TASL-2* Δ HP) and placed the blasticidin resistance cassette 800 bp downstream from the predicted *TASL* transcript (Fig. 7F, left). Following selection, wild-type and deletion chromosomes were detected by Southern blot (Fig. 7E, F). The *TASL-2* Δ TX chromosome could not fully replace the wild-type chromosome, but in contrast, *TASL-2* Δ UP did completely replace wild-type *TASL-2* (Fig. 7E). *TASL-2* Δ HP also completely replaced *TASL-2* (Fig. 7F), contrary to expectation if this distinctive feature of *TASL* transcripts gave the *TASL* ncRNAs an essential function. Surprisingly, *TASL-2* Δ UP and *TASL-2* Δ HP cells retained *TASL-2* ncRNA expression detectable by RT-PCR (data not shown).

TASL sRNAs are required for sexual reproduction

We proposed that the *TASL*-transcript hairpins are determinants for RDRC recruitment and sRNA biogenesis (Couvillion et al. 2009). To test this hypothesis, we assayed for sRNAs from *TASL-2* by Northern blot in *TASL-2* Δ HP cells. The most abundant *TASL* sRNAs are the phased, antisense-strand sRNAs carried by *Twis2p* (Couvillion et al. 2009). Oligonucleotide probe complementary to a *Twis2p*-bound *TASL-2* sRNA detected 23- to 24-nt sRNA in Δ *Twis8* cells but no sRNA in *TASL-2* Δ HP cells (Fig. 8A). This suggests that *TASL* transcript hairpins are indeed determinants of sRNA biogenesis.

We used *TASL-2* Δ HP cells of different mating types to investigate whether loss of *TASL* sRNAs has the same phenotype as loss of *Twis8p*. We mated two wild-type, two *TASL-2* Δ UP or *TASL-2* Δ HP strains, and also wild-type with *TASL-2* Δ UP or *TASL-2* Δ HP. All combinations formed cell pairs (data not shown). Paired cells were scored for completion of conjugation at 24 h post-mixing. Only the conjugation of two *TASL-2* Δ HP strains failed to produce exconjugants with the wild-type end-point of

two macronuclei and one micronucleus (Fig. 8B). The majority of paired *TASL-2* Δ HP cells arrested early in zygotic nuclei differentiation, paralleling the *Twis8p* loss-of-function phenotype. We conclude that both *Twis8p* and the sRNAs

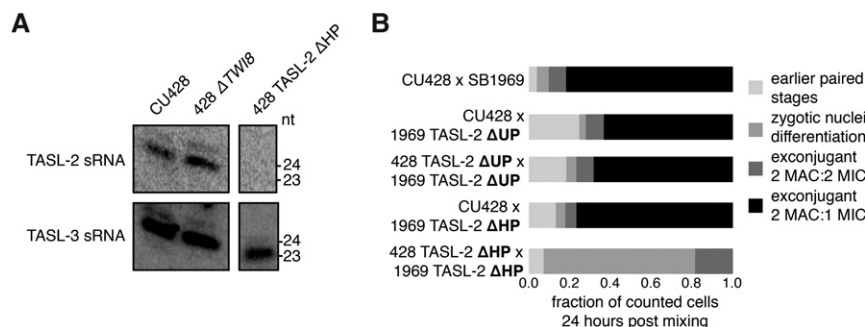


FIGURE 8. *TASL-2*-derived sRNAs are required for completion of conjugation. (A) Northern blots of sRNAs derived from *TASL-2* and *TASL-3* RNAs. Total RNA was collected from starved cultures of the listed strains and enriched for sRNAs by collecting the flow-through following centrifugal concentration. Enriched populations of sRNAs were resolved on a denaturing acrylamide gel and electrotransferred to a nylon membrane. *TASL-2* and *TASL-3* sRNAs were detected using oligonucleotide probes antisense to the single most abundant sRNA from each locus as determined from the data presented in Figure 5. All lanes are cropped from the same exposure of the same blot. (B) Mating progression of conjugating pairs bearing *TASL-2* modifications. Starving cultures of the listed cells were mixed and allowed to pair and progress through conjugation. Twenty-four hours after mixing, conjugating pairs were collected, fixed, and stained with DAPI to determine progression through conjugation. Progression was determined using the categories and criteria established in Figure 2. At least 100 cells were counted from each cross.

loaded to Twi8p are essential for sexual reproduction, but only for sexual reproduction. In contrast, the DNA loci that generate Twi8p-bound sRNAs are also essential for asexual growth. Furthermore, individual TASLs have nonredundant functions: Loss of sRNAs from even a single TASL imposed the complete failure of sexual reproduction, and decreased DNA copy number of even a single TASL was deleterious for asexual cell growth.

DISCUSSION

Like other organisms, *Tetrahymena* generates more piRNAs with transposon-derived than unique-locus sequence (Lee and Collins 2006; Iwasaki et al. 2015). However, specialization among *Tetrahymena* Piwi proteins allowed us to resolve their functions in transposon control from a distinct function at ncRNA loci. Our findings suggest the possibility of broader activities for eukaryotic Piwi proteins in general.

Biogenesis of *Tetrahymena* 23- to 24-nt sRNAs and RNPs

Tetrahymena Dcr2p cleaves RDRC-synthesized dsRNA to generate 23- to 24-nt sRNAs (Lee and Collins 2007). The transcripts degraded by Dcr2p are determined by three RDRCs, all sharing the catalytic subunit Rdr1p. Rdr1p assembles with the mutually exclusive, paralogous, polyuridine polymerase subunits Rdn1p and Rdn2p, and Rdn1p recruits the mutually exclusive, paralogous subunits Rdf1p and Rdf2p (Lee and Collins 2007; Lee et al. 2009). To account for the cellular specificities of Twi-protein sRNA loading, based on this study and previous work, we propose that the Rdf1 RDRC generates dsRNA loaded to Twi8p (Fig. 9A, right), possibly

in the macronucleus, while Rdf2 and Rdn2 RDRCs generate dsRNA loaded to Twi2p in the cytoplasm (Fig. 9A, center and left).

RDF1 knockout cells are viable for vegetative growth, but without parental Rdf1p, paired cells arrest in conjugation shortly after the generation of four zygotic nuclei (Lee et al. 2009). Knockout of *RDF1* did not eliminate any type of sRNA abundant enough to detect by Northern blot or bulk sRNA sequencing (Couvillion et al. 2009). These phenotypes match expectation for a deficiency of Twi8 RNP biogenesis. Therefore, we suggest that a pathway initiated by the Rdf1 RDRC loads Twi8p with sense and antisense strands of TASL primary transcripts (Fig. 9A, left). The small pool of Twi8p-bound sense and antisense pseudogene-derived sRNAs suggests that with low efficiency, the Rdf1 RDRC recognizes other primary transcripts including those of PGC family I, which could form hairpin structures in their degenerate repeat regions (Lee and Collins 2006). *RDF1* mRNA level is relatively low in early conjugation (Lee et al. 2009), paralleling levels of *TWI8* mRNA assayed in this work.

RDF2 knockout eliminates the abundant Twi2p-bound TASL sRNAs that are exclusively antisense to the primary transcript and phased (Couvillion et al. 2009). Thus, TASL transcripts not acted on by the Rdf1 RDRC are eliminated in the cytoplasm by the Rdf2 RDRC (Fig. 9A, center). *RDF2* knockout cells are viable for vegetative growth, and parental *RDF2* is not required for sexual reproduction (Lee et al. 2009). This is consistent with the lack of vegetative growth or sexual reproduction phenotype upon knockout of *TWI2* or the entire *TWI2-TWI6* locus in parental macronuclei (Couvillion et al. 2009) (additional data not shown). In the absence of the Rdf2 RDRC, aberrant transcripts that reach the cytoplasm are likely degraded by nonsense-mediated mRNA decay.

RDN2 knockout dramatically reduces Twi2p-bound pseudogene-derived sRNAs that are antisense to the primary transcript (Fig. 9A, right) (Couvillion et al. 2009). Some PGC transcripts may also be degraded by the Rdf2 RDRC (Fig. 9A, dashed arrow), based on more comprehensive PGC sRNA loss in cells with combined *RDN2* and *RDF2* knockout (Couvillion et al. 2009; Talsky and Collins 2012). *RDN2* mRNA level is maximal during sexual reproduction (Lee et al. 2009). Like *RDF1* and *TWI8*, *RDN2* macronuclear gene knockout cells are viable for vegetative growth but arrest in sexual reproduction after generation of four zygotic nuclei (Lee et al. 2009). Because *TWI2* is not essential for sexual reproduction, the critical function of the Rdf2 RDRC is likely to be transcript degradation. We suggest that the enormous

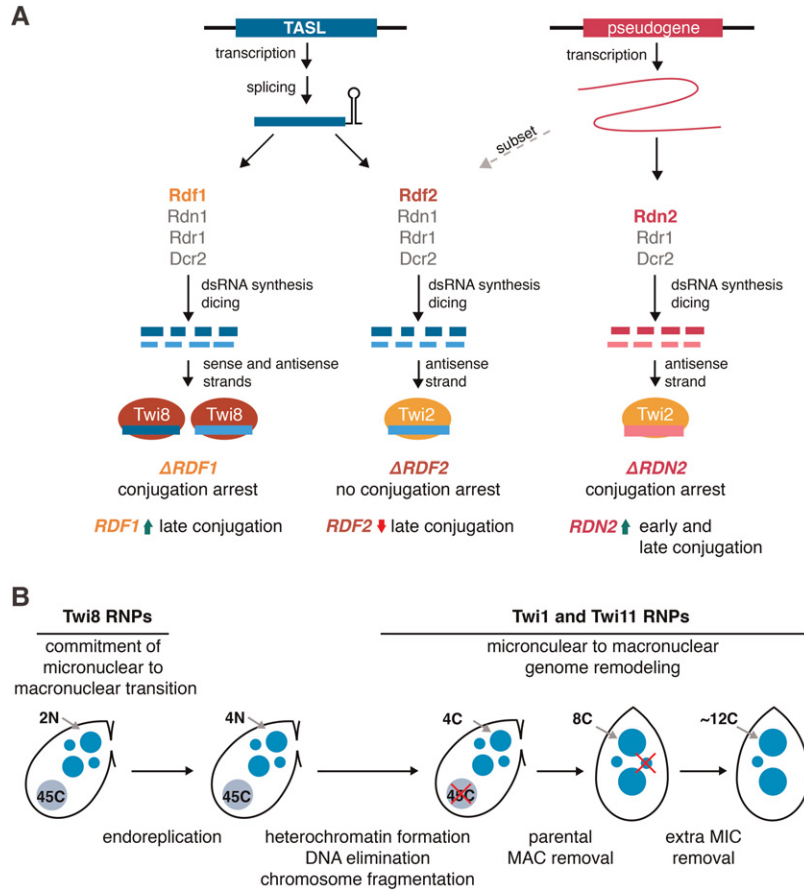


FIGURE 9. Proposed functions for Twi proteins in DNA amplification and elimination. (A) Model for the pathways of *Tetrahymena* 23- to 24-nt sRNA biogenesis and RNP assembly, as described in the Discussion. (B) Current model of possible *Tetrahymena* nuclear Piwi protein functions in sexual reproduction. N indicates germline chromosome copy number while C indicates macronuclear chromosome copy number. Which process(es) in early zygotic macronucleus differentiation are directly influenced by Twi8p and/or Twi8 RNPs is not yet known.

amount of nongenic transcription in early sexual reproduction obliges a clean-up system that the Rdf2 RDRC provides.

Functions of Twi8p and TASLs

The accumulation of TASL ncRNA transcripts increases in the absence of Twi8p, shown here, or the absence of the Rdf2 RDRC that loads Twi2p (Couvillion et al. 2009). Neither Twi8p nor Twi2p has the active site amino acids required for endonucleolytic cleavage activity (Liu et al. 2004; Couvillion et al. 2009). We suggest that the Twi8p- and Twi2p-dependent decrease in TASL transcript levels occurs largely due to sRNA-biogenesis: Rdf1 RDRC degradation of TASL transcripts loads Twi8p with balanced levels of sRNAs from each TASL, and any remaining TASL transcripts are degraded by Rdf2 RDRC loading of Twi2p (Fig. 9A). Additionally, if Twi8p RNPs have an effector function, they could decrease TASL transcript levels by cotranscriptional silencing (Guang et al. 2010). TASLs have genome neighborhoods devoid of nearby mRNA expression that are highly

atypical in the gene-dense macronuclear genome streamlined by DNA elimination (Coynne et al. 2008). It will be of future interest to determine how TASLs are transcribed, how sRNAs from each TASL come to be equally represented in the Twi8p-bound sRNA pool, and whether Twi8 RNPs act on TASLs or their transcripts through sRNA base-pairing.

The genetically essential role of Twi8p is in sexual reproduction. Mitotic divisions of the zygotic nucleus give rise to two future macronuclei per cell, which undergo rapid endoreplication to 4N (Fig. 9B; N indicates germline chromosome copy number while C indicates macronuclear chromosome copy number). Zygotic nuclei committed to macronuclear differentiation retain 4N content for several hours, ultimately undergoing DNA elimination and chromosome breakage followed by additional amplification (Fig. 9B). The conjugation arrest point for $\Delta TWI8$ cells is earlier than typical for cells mutant in heterochromatin assembly factors, because $\Delta TWI8$ pairs retain parental macronuclei instead of destroying them. Several lines of evidence support the conclusion that zygotic function of Twi8 RNPs is fulfilled by RNPs assembled in a parental macronucleus. First, *TWI8* mRNA expression is low in early conjugation from both parental and zygotic nuclei. Second, the 2'-O-methylation of Twi8p-bound sRNA 3' ends would decrease their biological turnover (Yu et al. 2005; Couvillion et al. 2009; Kurth and Mochizuki 2009). Third, the Twi8p-associated sRNA pool does not change in composition across the *Tetrahymena* lifecycle.

At the molecular level, one possible model for Twi8 RNP function is to license the micronucleus-length chromosomes for macronucleus differentiation. It is perhaps not a coincidence that all five micronuclear chromosomes have at least one of the seven TASLs. Micronuclei can lose entire chromosomes in vegetative growth, since phenotypic selection is only on macronucleus genome content. Twi8 RNPs could satisfy a checkpoint for germline genome integrity, or they could function directly to trigger a process such as endoreplication. A mechanism of ncRNA-targeted DNA replication initiation has been previously proposed for *Tetrahymena* rDNA chromosome amplification (Mohammad et al. 2007).

Unlike the germline genome, the somatic genome does not have a TASL on every chromosome. Thus it was surprising that even an individual TASL could not be substantially decreased in copy number in the macronucleus of cells in

vegetative growth. It seems unlikely that targeting of each TASL disrupted the function of an adjacent gene. Instead, the TASLs themselves appear to have an essential function. In retrospect, a mechanism that obliged retention of each TASL at high copy number would ensure that progeny of sexual reproduction inherit *Twis8* RNPs with sRNAs from every TASL. The growth-essential function of TASLs could be linked to the dsRNA binding protein *Drb2p*, which forms foci in macronuclei that could be at TASLs (Motl and Chalker 2011). Like TASLs, *Drb2p* is essential in vegetative growth and sexual reproduction (Motl and Chalker 2011). Loss of zygotic-nucleus *DRB2* arrested conjugation prior to DNA elimination and chromosome fragmentation but after paired cells separated to exconjugants. Morphologically this is a later arrest point than in Δ *TWIS8* cells, but the difference could be due to some *Drb2p* expression in early conjugation from the wild-type parental macronuclei of the germline-genome knockouts. We suggest that TASLs provide an opportunity to understand the biological function(s) of dsRNA in and beyond its roles in epigenetic inheritance.

MATERIALS AND METHODS

Strain construction and culture

Strain construction used established methods (Chalker 2012; Couvillion and Collins 2012). The N-terminal ZZF-*Twis8p* tag was introduced to the *TWIS8* locus with a *BSR2* cassette upstream of the predicted *TWIS8* promoter. The *TWIS8* knockout construct replaced exons 4 and 5 of *TWIS8* with a *BSR2* cassette. The initial TASL knockout constructs replaced the sequence with mapped sRNA reads and ~300-bp flanking sequence on both sides with either a *BSR2* or *NEO4* cassette. Additional TASL-2 deletion constructs replaced either the sequence to which sRNAs mapped (TASL-2 Δ TX) or 800 bp upstream of that sequence (TASL-2 Δ UP) with a *BSR2* cassette. The TASL-2 Δ HHP deletion construct removed the predicted hairpin sequence and inserted a *BSR2* cassette 800 bp downstream from the predicted hairpin. Each construct was introduced by biolistic transformation. Following transformation, maximal assortment to the recombinant chromosome was performed by growth in increasing concentrations of either blasticidin (for *BSR2* cassette) or paromomycin (for *NEO4* cassette). Individual cells were picked from maximal selection to establish clonal cell lines, which were returned to vegetative growth without selection. Cells were cultured for vegetative growth in 0.25% proteose peptone, 0.25% yeast extract, and 0.5% glucose supplemented with 10 μ M FeCl_2 at 30°C with shaking. For conjugation, cells were harvested in log phase of vegetative growth, transferred to starvation medium (10 mM Tris, pH 7.5), diluted to 200,000 cells/mL, and maintained at 30°C with shaking for between 16 and 24 h. To initiate conjugation, equal volumes of cells of two mating types were mixed and maintained in cultures <1 cm deep at 30°C without shaking.

Viability test for progeny of sexual reproduction

Cells were mated as above and transferred to 2% proteose peptone, 0.2% yeast extract, 0.2% glucose supplemented with 10 μ M FeCl_2 at

24 h post-mixing. After 24 h of growth at 30°C, cells were diluted 1:50 with either the same media or the same media supplemented with 15 μ g/mL cycloheximide. Only cells with a functional zygotic macronucleus will have a cycloheximide-resistant phenotype. Viability was scored after 3 d of culture in cycloheximide by counting live cells.

Nucleic acid detection

M-element PCR detection was performed as previously described (Mochizuki et al. 2002). Detection of the rDNA intermediate was performed as previously described (Pan and Blackburn 1981). For Northern blots of total cellular RNA, 20 μ g of TRIzol-purified RNA per sample was resolved on a formaldehyde agarose gel and transferred to nylon membrane. The ladder was trimmed from each membrane and visualized by methylene blue staining. For Southern blots, the DNA marker ladder was detected by ethidium bromide staining of the gel pre-transfer. Southern and Northern blot probes for mRNA detection were made by random-hexamer priming using a PCR product as template. For Northern blot detection of sRNAs, total RNA samples were enriched for sRNAs by collecting the flow-through from centrifugal concentration using a Millipore Microcon 10K spin column, ethanol precipitated, resolved by denaturing PAGE, transferred to nylon membrane, and detected with 5'-end-radiolabeled DNA oligonucleotide. To detect TASL knockouts in the amiconucleate strain BI3840, we used multiplex PCR. Twenty-five nanograms of genomic DNA was used in a 50 μ L reaction containing the following: 1 \times Q5 reaction buffer, 0.2 mM dNTPs, 0.5 units Q5 polymerase, and 0.25 μ M each of forward and reverse primers amplifying *TBPI* and the replaced region of the targeted TASL. Thirty-five cycles of amplification were performed with the following parameters: 30 sec denaturation at 95°C, 30 sec annealing at 61°C, and 30 sec extension at 72°C. PCR products were resolved on a 1.5% agarose gel.

Cell staining

Cells were fixed for immunofluorescence as previously described (Nelsen et al. 1994). Briefly, cells were harvested by centrifugation and resuspended in residual buffer. Tubes were transferred to ice and incubated for 2 min, at which point ice-cold 50% EtOH/0.2% Triton X-100 was added. Cells were fixed for 5 min on ice, washed with 1 \times TBS (20 mM Tris-HCl at pH 7.4, 150 mM NaCl)/0.1% bovine serum albumin twice, and incubated with primary antibody at 4°C overnight. Fixed cells were washed twice with 1 \times TBS/0.1% bovine serum albumin and incubated with secondary antibody for 2–3 h at room temperature. Cells were then washed twice more, resuspended in Vectashield (Vector Laboratories) supplemented with 25 ng/mL DAPI, and spread on slides for microscopy. Imaging was performed on a Zeiss Axioimager M1 widefield microscope with epi-illumination outfitted with a Hamamatsu C8484 camera. Staining for ZZF-*Twis8p* was performed with mouse anti-Flag antibody (1:100 dilution, Sigma F1804), staining of Pdd1p was performed with rabbit anti-Pdd1p antibody (1:250 dilution, Abcam ab5338), and staining of H3K9me3 was performed with rabbit anti-H3K9me3 antibody (1:500 dilution, Millipore 07-442). Either goat anti-rabbit (Life Technologies A-11008) or goat anti-mouse (Life Technologies A-11029) secondary antibodies labeled by Alexa Fluor 488 were used at a 1:1000 dilution.

For DAPI staining only, cells were fixed as previously described (Loidl et al. 2012). Briefly, cells were collected in centrifuge tubes then Triton X-100 and formaldehyde were added to final concentrations of 0.5% and 3.7%, respectively. Cells were fixed with rotation at room temperature for 30 min, centrifuged, and resuspended in 4% paraformaldehyde with 3.4% glucose. Cells were spread on a microscope slide and allowed to dry. Following rehydration in 1× PBS (10 mM sodium phosphate at pH 7.4, 150 mM NaCl, 0.5 mM MgCl₂), cells were stained with 25 ng/mL DAPI and imaged using the microscope described above.

RNP affinity purification, sRNA cloning, and bioinformatic analyses

Affinity purification of ZZF-Twis8 RNPs was performed as previously described (Couvillion and Collins 2012). Briefly, cells were collected by centrifugation (cells in vegetative growth were washed briefly into starvation medium) and lysed by the addition of 20 mM Tris at pH 8.0, 50 mM NaCl, 1 mM EDTA, 10% glycerol (T2EG50) supplemented with 0.2% Igepal CA-630, 0.1% Triton X-100, 5 mM DTT, 5 μM MG-132, 0.1 mM phenylmethanesulfonyl fluoride, and 1× Sigma Protease Inhibitor. Lysates were cleared by ultracentrifugation and incubated with IgG agarose, followed by agarose washing with T2EG50 + 0.1% Igepal + 1 mM DTT. Bound Twis8 RNPs were eluted by cleavage with Tobacco Etch Virus protease, and the eluate was incubated with Flag M2 antibody resin. Unbound protein was washed away with T2EG50 + 0.1% Igepal + 1 mM DTT. Bound Twis8 RNPs were eluted by triple-Flag peptide.

Purified sRNAs were collected by phenol–chloroform extraction followed by ethanol precipitation. Size selection of purified sRNAs was performed by denaturing PAGE. sRNAs were prepared for Solexa sequencing as previously described (Malone et al. 2012). Briefly, IDT miRNA cloning linker 1 (Modban) was ligated to the 3′ end of the purified sRNAs using T4 RNA ligase, truncated, K227Q (NEB). Ligations were incubated at 16°C overnight to minimize the impact of 3′-terminal 2′O-methylation on the sRNAs (Munafo and Robb 2010). Ligated sRNAs were purified by denaturing PAGE and collected by ethanol precipitation. The sequencing adaptor was ligated to sRNA 5′ ends by T4 RNA ligase 1, and the intended products were gel-purified and collected by ethanol precipitation. Doubly ligated sRNAs were reverse-transcribed with Superscript III, and NEBNext multiplex adaptors were added by PCR. Library insert size was checked by Bioanalyzer. Libraries were pooled and sequenced on an Illumina HiSeq 2000.

The six libraries yielded an average of 4.9 million reads (high: CU428-starving, 5.9 million; low: SB1969-conjugating, 2.8 million). Reads that passed the internal Solexa filters had their adaptor sequences trimmed using fastq-mcf (<https://expressionanalysis.github.io/ea-utils/>) with the following parameters: -L 30, -k 0. Reads were then mapped to the 2008 assembly of the *Tetrahymena* macronuclear genome (Coyne et al. 2008) and the v2 assembly of the *Tetrahymena* micronuclear genome (<http://www.broadinstitute.org/annotation/genome/Tetrahymena/MultiHome.html>) using Bowtie2 (Langmead and Salzberg 2012) and the following parameters: -phred33 -local -very-sensitive. Reads with internal mismatches were not considered for further analysis, but because many studies indicate that *Tetrahymena* sRNAs can have a 3′ end nucleotide that is not genome-encoded, we used a local alignment to allow for mismatches at read ends. Each of the libraries had low sequence

diversity, with an average of 76,435 unique reads per library (high: SB1969-starving: 103,207 unique reads; low: SB1969-conjugating: 26,915 unique reads). The vast majority of reads mapped to sequences retained in the macronucleus (99%–100%; high: CU428-vegetative, 99.7%; low: SB1969-conjugating, 98.9%).

Predicted hairpin sequences were identified using RNALfold from the ViennaRNA Package 2 (Lorenz et al. 2011) and the following parameters: -L 150 -noLP. This searched for local secondary structures within a 150-nt window and disallowed the formation of structures containing single base pairs.

DATA DEPOSITION

sRNA RNA-seq data sets are available at the NCBI Sequence Read Archive, BioProject accession ID PRJNA306046.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank Mary Couvillion, Suzanne Lee, Kyungah Hong, and Nicole Beier for strains, plasmids, and/or information; Ed Orias and Robert Coyne for sharing unpublished data; and Collins laboratory members for discussion and comments. This work was supported by National Institute of General Medical Sciences, National Institutes of Health F32-GM109673 (B.M.F.) and R01-GM054198 (K.C.). This work was performed in part at the University of California at Berkeley Vincent J. Coates Genomics Sequencing Laboratory, supported by National Institutes of Health S10 Instrumentation grants S10-RR029668 and S10-RR027303I, and the University of California, Berkeley College of Natural Resources Biological Imaging Facility.

Received November 16, 2016; accepted December 29, 2016.

REFERENCES

- Austerberry CF, Allis CD, Yao MC. 1984. Specific DNA rearrangements in synchronously developing nuclei of *Tetrahymena*. *Proc Natl Acad Sci* **81**: 7383–7387.
- Castel SE, Martienssen RA. 2013. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat Rev Genet* **14**: 100–112.
- Cerutti H, Casas-Mollano JA. 2006. On the origin and functions of RNA-mediated silencing: from protists to man. *Curr Genet* **50**: 81–99.
- Chalker DL. 2012. Transformation and strain engineering of *Tetrahymena*. *Methods Cell Biol* **109**: 327–345.
- Chalker DL, Meyer E, Mochizuki K. 2013. Epigenetics of ciliates. *Cold Spring Harb Perspect Biol* **5**: a017764.
- Cheng CY, Vogt A, Mochizuki K, Yao MC. 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**: 1753–1762.
- Couvillion MT, Collins K. 2012. Biochemical approaches including the design and use of strains expressing epitope-tagged proteins. *Methods Cell Biol* **109**: 347–355.
- Couvillion MT, Lee SR, Hogstad B, Malone CD, Tonkin LA, Sachidanandam R, Hannon GJ, Collins K. 2009. Sequence, biogenesis,

- and function of diverse small RNA classes bound to the Piwi-family proteins of *Tetrahymena thermophila*. *Genes Dev* **23**: 2016–2032.
- Couvillion MT, Sachidanandam R, Collins K. 2010. A growth-essential *Tetrahymena* Piwi protein carries tRNA fragment cargo. *Genes Dev* **24**: 2742–2747.
- Couvillion MT, Bounova G, Purdom E, Speed TP, Collins K. 2012. A *Tetrahymena* Piwi bound to mature tRNA 3' fragments activates the exonuclease Xrn2 for RNA processing in the nucleus. *Mol Cell* **48**: 509–520.
- Coyne RS, Thiagarajan M, Jones KM, Wortman J, Tallon LJ, Haas BJ, Cassidy-Hanley DM, Wiley EA, Smith JJ, Collins K, et al. 2008. Refined annotation and assembly of the *Tetrahymena thermophila* genome sequence through EST analysis, comparative genomic hybridization, and targeted gap closure. *BMC Genomics* **9**: 562.
- Cranert S, Heyse S, Linger BR, Lescasse R, Price C. 2014. *Tetrahymena* Pot2 is a developmentally regulated paralog of Pot1 that localizes to chromosome breakage sites but not to telomeres. *Eukaryot Cell* **13**: 1519–1529.
- Eisen JA, Coyne RS, Wu M, Wu D, Thiagarajan M, Wortman JR, Badger JH, Ren Q, Amedeo P, Jones KM, et al. 2006. Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. *PLoS Biol* **4**: e286.
- Fass JN, Joshi NA, Couvillion MT, Bowen J, Gorovsky MA, Hamilton EP, Orias E, Hong K, Coyne RS, Eisen JA, et al. 2011. Genome-scale analysis of programmed DNA elimination sites in *Tetrahymena thermophila*. *G3* **1**: 515–522.
- Gebert D, Rosenkranz D. 2015. RNA-based regulation of transposon expression. *Wiley Interdiscip Rev RNA* **6**: 687–708.
- Guang S, Bochner AF, Burkhardt KB, Burton N, Pavelec DM, Kennedy S. 2010. Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. *Nature* **465**: 1097–1101.
- Holoch D, Moazed D. 2015. RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet* **16**: 71–84.
- Ipsaro JJ, Joshua-Tor L. 2015. From guide to target: molecular insights into eukaryotic RNA-interference machinery. *Nat Struct Mol Biol* **22**: 20–28.
- Iwasaki YW, Siomi MC, Siomi H. 2015. PIWI-interacting RNA: its biogenesis and functions. *Annu Rev Biochem* **84**: 405–433.
- Jonas S, Izaurralde E. 2015. Towards a molecular understanding of microRNA-mediated gene silencing. *Nat Rev Genet* **16**: 421–433.
- Kataoka K, Noto T, Mochizuki K. 2016. Phosphorylation of an HP1-like protein is a prerequisite for heterochromatin body formation in *Tetrahymena* DNA elimination. *Proc Natl Acad Sci* **113**: 9027–9032.
- Kobayashi H, Tomari Y. 2016. RISC assembly: coordination between small RNAs and Argonaute proteins. *Biochim Biophys Acta* **1859**: 71–81.
- Kuhn CD, Joshua-Tor L. 2013. Eukaryotic Argonautes come into focus. *Trends Biochem Sci* **38**: 263–271.
- Kurth HM, Mochizuki K. 2009. 2'-O-methylation stabilizes Piwi-associated small RNAs and ensures DNA elimination in *Tetrahymena*. *RNA* **15**: 675–685.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.
- Le Thomas A, Toth KF, Aravin AA. 2014. To be or not to be a piRNA: genomic origin and processing of piRNAs. *Genome Biol* **15**: 204.
- Lee SR, Collins K. 2006. Two classes of endogenous small RNAs in *Tetrahymena thermophila*. *Genes Dev* **20**: 28–33.
- Lee SR, Collins K. 2007. Physical and functional coupling of RNA-dependent RNA polymerase and Dicer in the biogenesis of endogenous siRNAs. *Nat Struct Mol Biol* **14**: 604–610.
- Lee SR, Talsky KB, Collins K. 2009. A single RNA-dependent RNA polymerase assembles with mutually exclusive nucleotidyl transferase subunits to direct different pathways of small RNA biogenesis. *RNA* **15**: 1363–1374.
- Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song JJ, Hammond SM, Joshua-Tor L, Hannon GJ. 2004. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**: 1437–1441.
- Liu Y, Taverna SD, Muratore TL, Shabanowitz J, Hunt DF, Allis CD. 2007. RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in *Tetrahymena*. *Genes Dev* **21**: 1530–1545.
- Loidl J, Lukaszewicz A, Howard-Till RA, Koestler T. 2012. The *Tetrahymena* meiotic chromosome bouquet is organized by centromeres and promotes interhomolog recombination. *J Cell Sci* **125**: 5873–5880.
- Lorenz R, Bernhart SH, Honer Zu Siederdisen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA Package 2.0. *Algorithms Mol Biol* **6**: 26.
- Luteijn MJ, Ketting RF. 2013. PIWI-interacting RNAs: from generation to transgenerational epigenetics. *Nat Rev Genet* **14**: 523–534.
- Madireddi MT, Coyne RS, Smothers JF, Mickey KM, Yao MC, Allis CD. 1996. Pdd1p, a novel chromodomain-containing protein, links heterochromatin assembly and DNA elimination in *Tetrahymena*. *Cell* **87**: 75–84.
- 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**: 9151–9164.
- Malone C, Brennecke J, Czech B, Aravin A, Hannon GJ. 2012. Preparation of small RNA libraries for high-throughput sequencing. *Cold Spring Harb Protoc* **2012**: 1067–1077.
- Meister G. 2013. Argonaute proteins: functional insights and emerging roles. *Nat Rev Genet* **14**: 447–459.
- 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**: 77–89.
- 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**: 689–699.
- Mohammad MM, Donti TR, Sebastian Yakisich J, Smith AG, Kapler GM. 2007. *Tetrahymena* ORC contains a ribosomal RNA fragment that participates in rDNA origin recognition. *EMBO J* **26**: 5048–5060.
- Motl JA, Chalker DL. 2011. Zygotic expression of the double-stranded RNA binding motif protein Drb2p is required for DNA elimination in the ciliate *Tetrahymena thermophila*. *Eukaryot Cell* **10**: 1648–1659.
- Munafo DB, Robb GB. 2010. Optimization of enzymatic reaction conditions for generating representative pools of cDNA from small RNA. *RNA* **16**: 2537–2552.
- Nakanishi K. 2016. Anatomy of RISC: how do small RNAs and chaperones activate Argonaute proteins? *Wiley Interdiscip Rev RNA* **7**: 637–660.
- Nelsen EM, Williams NE, Yi H, Knaak J, Frankel J. 1994. “Fenestrin” and conjugation in *Tetrahymena thermophila*. *J Eukaryot Microbiol* **41**: 483–495.
- Noto T, Kataoka K, Suhren JH, Hayashi A, Woolcock KJ, Gorovsky MA, Mochizuki K. 2015. Small-RNA-mediated genome-wide trans-recognition network in *Tetrahymena* DNA elimination. *Mol Cell* **59**: 229–242.
- Pan WC, Blackburn EH. 1981. Single extrachromosomal ribosomal RNA gene copies are synthesized during amplification of the rDNA in *Tetrahymena*. *Cell* **23**: 459–466.
- Ross RJ, Weiner MM, Lin H. 2014. PIWI proteins and PIWI-interacting RNAs in the soma. *Nature* **505**: 353–359.
- Ruehle MD, Orias E, Pearson CG. 2016. *Tetrahymena* as a unicellular model eukaryote: genetic and genomic tools. *Genetics* **203**: 649–665.
- Sarkar A, Volff JN, Vaury C. 2016. piRNAs and their diverse roles: a transposable element-driven tactic for gene regulation? *FASEB J*. doi: 10.1096/fj.201600637RR.
- 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**: 1729–1742.
- Seto AG, Kingston RE, Lau NC. 2007. The coming of age for Piwi proteins. *Mol Cell* **26**: 603–609.
- Talsky KB, Collins K. 2010. Initiation by a eukaryotic RNA-dependent RNA polymerase requires looping of the template end and is influenced by the template-tailing activity of an associated uridylyltransferase. *J Biol Chem* **285**: 27614–27623.

- Talsky KB, Collins K. 2012. Strand-asymmetric endogenous *Tetrahymena* small RNA production requires a previously uncharacterized uridylyl-transferase protein partner. *RNA* **18**: 1553–1562.
- Taverna SD, Coyne RS, Allis CD. 2002. Methylation of histone h3 at lysine 9 targets programmed DNA elimination in *Tetrahymena*. *Cell* **110**: 701–711.
- Weick EM, Miska EA. 2014. piRNAs: from biogenesis to function. *Development* **141**: 3458–3471.
- Wenkert D, Allis CD. 1984. Timing of the appearance of macronuclear-specific histone variant hv1 and gene expression in developing new macronuclei of *Tetrahymena thermophila*. *J Cell Biol* **98**: 2107–2117.
- Yao MC, Choi J, Yokoyama S, Austerberry CF, Yao CH. 1984. DNA elimination in *Tetrahymena*: a developmental process involving extensive breakage and rejoining of DNA at defined sites. *Cell* **36**: 433–440.
- Yao MC, Zheng K, Yao CH. 1987. A conserved nucleotide sequence at the sites of developmentally regulated chromosomal breakage in *Tetrahymena*. *Cell* **48**: 779–788.
- Yao MC, Chao JL, Cheng CY. 2014. Programmed genome rearrangements in *Tetrahymena*. *Microbiol Spectr* **2**. doi: 10.1128/microbiol-spec.MDNA3-0012-2014.
- Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, Steward R, Chen X. 2005. Methylation as a crucial step in plant microRNA biogenesis. *Science* **307**: 932–935.