universität
wien

DISSERTATION

Titel der Dissertation

Analysis of COnjugation Induced (COI) genes in *Tetrahymena thermophila*

Verfasserin

Lucia Aronica

angestrebter akademischer Grad

Doctor of Philosophy (PhD)

Wien, am 4. Januar 2010

Matrikel-Nummer:	0606551
Dissertationsgebiet lt. Studienblatt:	Molekulare Biologie
Betreuer:	Dr. Kazufumi Mochizuki

Table of contents

Introduction

Results

Discussion

Material and methods

Zusammenfassung

In *Tetrahymena thermophila* kommt es zu programmierter Eliminierung bestimmter DNA Sequenzen waehrend der sexuellen Reproduktion (Konjugation).

Ziel dieser Arbeit war es, Gene zu identifizieren, die an der DNA Eliminierung beteiligt sind. 22 konjugations-spezifische Gene (COI Gene) konnten als potenzielle Kandidaten identifiziert werden.

Im Verlauf dieser Arbeit wurde *COI12*, das eine spezifische Expression waehrend der Konjugation zeigt und notwendig ist fuer die Produktion lebensfaehiger Nachkommen, genauer analysiert. *COI12* ist essentiell fuer die Eliminierung von DNA und agiert nach dem Export der scnRNA aus dem Mic in das Zytoplasma und vor dem Beladen des Argonaut Proteins Twi1p mit scnRNA. Ausserdem ist eine Primer-abhaengige RNA-abhaengige RNA Polymerase- (RdRP) Aktivitaet mit dem Protein Coi12p assoziiert. Coi12p interagiert mit zwei Proteinen, die Homologie zu bekannten RdRP Faktoren aufzeigen, Rdn3p and Rdf3p, und mit dem Dicer Protein Dcr1p. Moeglicherweise wird durch den mit Coi12p assoziierten RDRC dsRNA produziert, die von Dcr1p in sekundaere scnRNA prozessiert wird. Diese neue Klasse von scnRNA koennte spezifisch in Twi1p geladen werden und als epigenetischer Faktor im Prozess der DNA Eliminierung dienen.

Abstract

Tetrahymena thermophila, like other ciliates, carries out programmed DNA elimination during sexual reproduction (conjugation).

Aim of this project was to identify novel genes involved in DNA elimination. Twenty-two conjugation specific genes (COI genes) have been identified as candidate gene with a possible role in DNA elimination.

In this study, I identified a COnjugation Induced (*COI*) gene, *COI12*, which is specifically expressed during conjugation and required for producing viable sexual progeny. I showed that *COI12* is essential for DNA elimination and acts downstream scnRNA export from the Mic to the cytoplasm and upstream scnRNA loading into the Argonaute Twi1p. Moreover, Coi12p is associated to a primer-dependent RNA-dependent RNA polymerase (RdRP) activity and interacts with two proteins related to RdRP factors, Rdn3p and Rdf3p, as well as with the Dicer Dcr1p.

Possibly, Coi12p-associated RDRC produces dsRNA that are processed by Dcr1p into a second class of scnRNA, secondary scnRNA, which is specifically loaded into Twi1p and act as epigenetic messenger in DNA elimination.

Introduction

1. Genomes in motion: DNA rearrangements in development and evolution

Genomes are not static. They undergo complex modifications driving cell development and evolution. Therefore, even in the same organism, different cells may have different genomes. Programmed DNA rearrangements are involved in genome remodeling in many organisms. They include processes that play a major role in cell differentiation - such as amplifications, deletions and re-arrangements of genes and gene segments.

DNA diminution is a fascinating type of programmed DNA rearrangement occurring at many genomic locations (Yao & Chao, 2005). DNA diminution includes chromosome elimination and chromatin diminution. In chromosome elimination entire chromosomes are removed during development, as observed in many sciarid flies (Goday & Esteban, 2001). In chromatin diminution individual portions of chromosomes are selected for removal. Chromatin diminution was observed for the first time in the nematode Ascaris more than a century ago. More recently, it has been implicated also in some species of crustaceans Cyclops (Beermann, 1977), in fish (Kubota et al., 1993) and in most ciliated protozoa (Prescott, 1994).

Most of the current knowledge on chromatin diminution and its molecular mechanism comes from studies in ciliated protozoa in the past thirty years. These investigations have uncovered an impressing variety of DNA remodeling mechanisms including amplification, deletion, elimination, inversion, deshuffling and dimerization of specific DNA sequences, as well as chromosome fragmentation and new telomere formation (Yao & Chao, 2005).

A remarkable example of chromatin diminution is DNA elimination in the ciliate *Tetrahymena*, a process which shows common features with both RNA-interference and heterochromatin formation in other organisms.

2. DNA rearrangements in *Tetrahymena thermophila*

2.1 *Tetrahymena thermophila*: a single cell with two different genomes

Like most ciliated protozoa, *Tetrahymena* shows a curious genetic feature referred as nuclear dimorphism: in a single cell there are two structurally and functionally different nuclei: the somatic macronucleus (Mac) and the germline micronucleus (Mic). Mac and Mic display their different cellular functions during the two main stages of *Tetrahymena* life cycle: vegetative growth and sexual reproduction, also named conjugation (Fig. 1). During vegetative growth the Mic is transcriptionally silent and the Mac governs the whole gene expression and thus the cell's phenotype. Conjugation takes place between starved cells of different mating types (Collins & Gorovsky, 2005). During conjugation (Fig. 2) the germline Mic is responsible for the transmission of the genome to the progeny, while the somatic Mac is destroyed. In this process, the Mic undergoes meiosis to form two haploid pronuclei, one of which is reciprocally exchanged between the two conjugating cells. The migratory and stationary pronuclei then fuse to create a diploid zygotic nucleus that divides mitotically twice to produce the next generation of new Macs and Mics. Then, paired cells separate, one of the two new Mics and the parental Mac are destroyed and, if refed, the exconjugants (progeny) resume vegetative growth (Allis-Jenuwein-Reinberg, 2006).

Mac and Mic differ also in their genetic asset and in the mechanism of nuclear division. Mics are diploid and they divide by typical mitosis. In contrast, Macs are ~45-ploid and divide by an unusual process of amitosis, in which no apparent chromosome condensation or spindle formation is involved. The risk of lethal gene loss is likely prevented by the high ploidy of Macs.

The Mac and Mic of *Tetrahymena* have been used as a model for active versus silent chromatin in epigenetic research (Fig. 3), which enabled the identification of important chromatin components and modifiers. For example, by comparing chromatin proteins in the Mac and Mic some of the first histone variants were identified. The observation that the histone variants hv1 and hv2 (corresponding to H2A.Z and H3.3 in other eukaryotes) were present only in the Mac, provided a first indication for a role of these proteins in maintaining active transcriptional activity (Hayashi et al., 1984; Allis et al., 1986). Moreover, new chromatin regulators, such as the first type A (nuclear) histone acetyltransferase (HAT) were identified in *Tetrahymena* by exploiting the presence of hyperacetylated histones only within the active chromatin of the Mac (Allis et al., 1985; Brownell et al., 1996). The finding that the *Tetrahymena* HAT was a homologue of the yeast transcription factor GCN5 established that transcriptional regulators may act as enzyme and modify chromatin (Brownell et al., 1996).

2.2 Overview of DNA rearrangements in *Tetrahymena thermophyla*

During sexual reproduction of *Tetrahymena,* the germline Mic produces the next generation of both new Mic and new Mac. This process involves four main steps: 1) DNA elimination ~15% of the Mic germline genome in the new somatic Mac, 2) chromosome breakage, 3) endoreplication and 4) selective amplification. The final genetic asset displays 45 copies of ~275 chromosomes and 9000 copies of a unique palindromic chromosome encoding ribosomal RNA (Collins & Gorovsky, 2005).

DNA elimination involves the specific removal of internal regions of the chromosome, without formation of free chromosome ends (Fig. 4). In *Tetrahymena*, it takes place at ~6000 Mic specific sequences referred as internal eliminated sequences (IESs) and it is followed by ligation of the flanking Mac destined sequences (MDSs). IESs vary from 0.5 to > 20kb in length and together represent about 15% of the genome (Yao & Gorovsky, 1974; Callahan et al., 1984). They have not been found within protein-coding sequences, although some are imbedded in introns. Most IESs are repeated in the Mic genome and some of them are related to trasposons such as Tlr-1 (Patil et al., 1997), Tel1 (Cherry & Blackburn, 1985) and REP (Fillingham et al., 2004) elements. Moreover, IES deletion intermediates show 4 bp protruding at the 5' ends (Saveliev & Cox, 1996) and thus resemble trans-esterification products of some transposons, including Tn7. Therefore, it has been proposed that IESs are remnants of transposons. IES excision is very precise and occurs reproducibly at specific sites or at a limited number of alternative sites (Yao & Chao, 2005). Nevertheless, so far no common consensus sequence has been identified. Certain common features between IESs may provide some hint about the mechanism of their deletion: the presence of short (1–8 bp) terminal direct repeats in most of them and the existence of *cis*-acting sequences that may control the specificity of DNA deletion. These cis-acting elements include a set of internal promoting sequences that determine the region to be deleted and a pair of flanking sequences (Godiska & Yao, 1990; Godiska et al., 1993) that specify the location of deletion boundaries. However, these regulatory elements vary markedly between different IESs and their mechanism of action is still elusive.

Chromosome breakage takes place after or in parallel to DNA elimination at roughly 200 specific sites in the genome and generates ~275 Mac chromosomes from the five chromosomes in the Mic (haploid) genome (Mochizuki & Gorovsky, 2004b). Unlike IESs, the chromosome breakage sites share a common 15-bp consensus sequence called chromosome breakage sequence (CBS) (King & Yao, 1982; Yao et al., 1990), which is necessary and sufficient to promote chromosome breakage and may represent a recognition site for unknown effector proteins. After chromosome breakage, small (<50 bp) sequences around CBSs are deleted and telomeres are added (Yao, 1981; King & Yao, 1982; Yao et al., 1990).

3 DNA elimination in *Tetrahymena* involves small RNA directed heterochromatin formation

3.1 DNA elimination involves heterochromatin formation

In *Tetrahymena*, eliminating IESs share similarity with heterochromatin. In many eukaryotes heterochromatin contains repeated sequences and its formation involves histone hypoacetylation, methylation of histone H3 at lysine 9 (H3K9) and 27 (H3K27) and accumulation of chromodomain proteins on the methylated histones. Similarly, most IESs are repeated sequences, their elimination is abolished by inhibiting histone deacetylation (Duharcourt & Yao, 2002) and is targeted by H3K9 (di-, tri-)

and H3K27 (tri) methylation catalyzed by histone methyltransferase Ezl1p (Taverna et al., 2002; Liu et al., 2007). The histone H3 methylations and IES eliminations are impaired in *EZL1* KO (Liu et al., 2007). Moreover, the chromodomain proteins Programmed DNA Deletion 1 and 3 (Pdd1p and Pdd3p), which is related to the chromodomain proteins HP1 of *animals* and SWI6 of *S. pombe,* colocalize and are associated with IESs (Nikiforov et al., 2000; Taverna et al., 2002). In vitro, Pdd1p binds H3K9me and H3K27me, while Pdd3p binds specifically H3K9me (Taverna et al., 2002; Liu et al., 2007). Pdd1p is required for IES elimination (Madireddi et al., 1996; Coyne et al., 1999) and its ectopic tethering to some locus is sufficient to promote its deletion (Taverna et al., 2002). Moreover, in many eukaryotes heterochromatin is assembled in organized structures, preferentially localized at the nucleus periphery. Similarly, eliminated IESs get clustered into nuclear foci in the new developing macronuclei (Madireddi et al., 1996; Smothers et al., 1997). These loci are believed to constitute DNA elimination structures and are referred as "dumposome", since they serve as selective garbage disposal for useless DNA.

Therefore, heterochromatin formation is involved in and required for DNA elimination in *Tetrahymena*. Heterochromatin has been related to DNA elimination also in other organisms, such as in Ascaris and Cyclops (Beerman S. 1977, Chromosoma. Tobler H. 1986. New York: Springer-Verlag Press).

3.2 DNA elimination is directed by small RNAs

3.2.1 RNA-interference related pathways: an overview

Recent evidence indicates that DNA-elimination in *Tetrahymena thermophila* is directed by small (~28 nucleotides) RNAs, by a mechanism sharing common features with RNA-interference in other creatures. Here, I first provide an overview of RNAi-related pathways.

The possibility that small RNA could have a role in gene regulation was suggested over 40 years ago (Jacob & Monod, 1961). In the last decade, studies in *C. elegans*, *S. pombe*, *Arabidopsis thaliana, Tetrahymena thermophila* and other model organism have established small (~20–30 nucleotide) RNAs as sequence-specific regulators of genes and their RNA products. RNA interference (RNAi) was originally described in *C. elegans* as a silencing mechanism where exogenous double-stranded RNAs (dsRNAs) could shut down the expression of their homologous sequences. This mechanism has been later related to other conserved pathways, all associated with small RNAs that target homologous sequences for inactivation by a variety of mechanisms. These pathways include post-transcriptional gene silencing (PTGS) and chromatin-dependent gene silencing (CDGS)- in which small RNA complexes are recruited on nascent transcripts (Moazed, 2009). CDGS include both transcriptional and co-transcriptional gene silencing (TGS and CTGS). The latter is mediated by RNA decay instead by transcription shut-off (Buhler et al., 2008). These RNA silencing pathways play key roles in gene regulation and genome stability- through stable heterochromatin formation at centromeres and telomeres- and in defense against transposons and RNA viruses- through degradation of RNA transcripts (Plasterk, 2002; Li & Ding, 2005).

RNA silencing pathways can be classified on the base of the effector small RNA molecules used. Three major classes of small RNA with different origin and cellular localization have been extensively characterized.

MicroRNAs (miRNAs) are ~21–25 nucleotides (nt) encoded in the genome. They are produced from hairpin precursors by the two ribonuclease III (RNase III) enzymes Drosha and Dicer (Fig. 5B). RNase III enzymes are ribonucleases specific for dsRNAs and typically generating products with characteristic two-nucleotide 3' overhangs. The hairpin precursors are first processed to pre-miRNAs by Drosha, which is mainly located in the nucleus, and the pre-miRNAs are eventually processed into shorter miRNA by Dicer in the cytoplasm. Most miRNAs have been shown to act in PTGS in the cytoplasm by mRNA degradation or translational arrest (Filipowicz et al., 2005). However, some plant miRNAs may direct DNA methylation in the nucleus (Bao et al., 2004) and promoter-directed human miRNAs were shown to be involved in TGS and induction of chromatin silencing (Gonzalez et al., 2008; Kim et al., 2008). Absent in fungi and other protozoa, miRNA pathways are conserved in plants and animals, although plants apparently lack Drosha enzyme.

The second class of small RNA, small interfering RNA (siRNAs), is also ~21–25 nucleotides in length and produced from long dsRNA by Dicer enzymes (Fig 5, A and C). These dsRNA originate from a variety of single-stranded RNA (ssRNA) precursors, including sense and antisense RNA transcripts arising from convergent promoters, hairpin RNAs transcribed from inverted repeat regions (Chung et al., 2008; Ghildiyal et al., 2008; Kawamura et al., 2008), and dsRNA products of RNA-dependent RNA polymerase (RdRP) enzymes. These enzymes recognize free 3’ ends of non-coding transcripts or aberrant RNAs that lack processing signals. siRNA pathways are conserved in a broad range of eukaryotes, from fungi to mammals.

Both miRNAs and siRNAs are loaded into two types of effector complexes: the cytoplasmic RNA-induced silencing complex (RISC) and the nuclear RNA-induced transcriptional silencing complex (RITS) (Hamilton & Baulcombe, 1999; Hammond et al., 2000; Zamore et al., 2000; Bernstein et al., 2001; Verdel et al., 2004). The small RNAs are bound by a highly conserved protein named Argonaute containing PAZ and PIWI domains. Although Dicer-processed si- and miRNAs are double-stranded, only one of their two strands , named the guide strand, is retained in both complexes, while the other strand, named the passenger strand is removed. Some Argonaute proteins not only bind small RNAs but also remove the passenger strand through their endoribonuclease (Slicer) activities residing in their conserved RNaseH-like PIWI domain (Okamura et al., 2004; Song & Joshua-Tor, 2006; Buker et al., 2007). However, other Argonaute proteins do not have Slicer activity and the mechanism by which they remove the passenger strands is not clear. Finally, the guide RNA-Argonaute complexs target their homologous sequences by base-pairing interactions. The target RNAs are cleaved by the Slicer activity of Argonaute proteins or are subjected to translational suppression.

The Argonaute-family can be divided phylogenetically into the AGO and PIWI clades (Carmell et al., 2002). The third class of small RNAs, the PIWI-interacting RNAs (piRNAs), associates with PIWI-clade Argonaute proteins. piRNAs vary in size from 24 to 31 nucleotides and derive from a variety of sequences, including repetitive DNA and transposons. They are produced and amplified by a “ping-pong” mechanism in which slicing of an antisense piRNA by a PIWI-Argonaute generates a sense piRNA which is in turn cleaved to produce a new antisense piRNA (Fig. 6) (Brennecke et al., 2007). piRNAs play key roles in the control of transposons in the germline of *D. melanogaster* and mammals (Aravin et al., 2007) as well as in repeat-induced and heterochromatic gene silencing in *D. melanogaster* somatic cells (Deshpande et al., 2005; Haynes et al., 2006). piRNAs have been identified in *D. melanogaster*, *C. elegans* and mammals, but not in fungi and plants .

3.2.2 The RNA link: first evidence and conclusive studies

Early in new Mac development, the IESs become enriched in H3K9 and H3K27 methylation, while the MDSs acquire modifications typical of euchromatin. A mechanism must exist to target specifically heterochromatin modifications to the IESs. However, in contrast to BESs, IESs share no common consensus sequence being able to act as a recognition site. Therefore, one could wonder: how are IESs recognized in absence of any genetic mark? Several evidences strongly suggest that small RNA molecules are involved in this process (Mochizuki et al., 2002; Yao et al., 2003).

First, Chalker and Yao (Chalker & Yao, 1996) demonstrated the existence of a sequence-specific communication between Mac and Mic, which could be mediated by a nucleic acid (Fig 6). Introduction of an IES into the Mac was shown to specifically inhibit elimination of the same IES in the new developing Mac. This curious finding suggested that the Mac controls epigenetically DNA elimination in the new Mac. However, it was still unclear which molecule could act as epigenetic messenger sent from the old Mac to the new Mac to target here IES elimination. The finding that IESs are bi-directionally transcribed during conjugation (Martindale et al., 1985; Chalker & Yao, 2001) provided a first indication that RNAs homologous to IESs could play such role. A breakthrough discovery came with the identification of an Argonaute ortholog, *TWI1*, which was shown to be required for IES elimination (Mochizuki et al., 2002). Since Argonaute proteins are conserved core components of RNAi-machineries, it was expected that small RNA were involved in DNA elimination. Indeed, it was found that small RNAs (~28 nucleotides long) homologous to IESs accumulate during conjugation. Moreover, disruption of *TWI1* destabilized these small RNA and impaired H3K9 methylation. These findings, together with those coming from a study in *S. pombe* (Volpe et al., 2002), brought together two areas of gene regulation previously thought to be unrelated: heterochromatin and RNAi. Indeed, it was shown that the bidirectional Mic transcripts are processed into the ~28nt RNAs by a Dicer ribonuclease, Dcl1p (Malone et al., 2005; Mochizuki & Gorovsky, 2005). The ~28 nucleotides (nt) RNA show 5' phosphate and 3' hydroxyl ends typically found in products of Dicer-like enzymes. Dcl1p is one of three Dicer-like proteins in *Tetrahymena*, together with Dcr1p and Dcr2p. Among these, only Dcl1p is specifically expressed during conjugation and required for the production of the ~28nt RNAs. Dcl1p localizes to the premeiotic micronuclei, indicating that small RNA production occurs at early stages of conjugation in the germline Mic. Disruption of *DCL1* abolishes small RNA production and IESs elimination (Malone et al., 2005; Mochizuki & Gorovsky, 2005).

3.2.3 The scnRNA model

It is believed that although the ~28 nt RNAs are produced from the entire micronuclear genome, including both IESs and MDSs, they are selected for IES homology in mid stages of conjugation. Indeed it has been observed that IES-specific small RNAs become enriched from mid to late stages of conjugation (Mochizuki & Gorovsky, 2004a) and most of the RNAs cloned at late conjugation are complementary to IES sequences (Lee & Collins, 2006).

How is the selection of the ~28 nt RNAs achieved? The old Mac must have a role in this process, based on the finding that introduction of an IES into the old Mac inhibits its excision from the new Mac (Chalker & Yao, 1996). Moreover, after being processed in the Mic by Dicer at early conjugation, the ~28 nt RNAs make a curious nuclear journey, with a stop-off point in the old Mac at mid stages and with final destination in the new Mac at late stages- as indicated by the cellular localization of their carrier protein, the Argonaute Twi1p. Based on these observations, the scnRNA model has been proposed (Mochizuki et al., 2002; Mochizuki & Gorovsky, 2004a) in which IES-specific RNAs are selected through a genome comparison ("scanning") between Mic and Mac and travel eventually to the new Mac to target epigenetically IES recognition and elimination. In this model, the ~28 nt RNAs, renamed as scnRNAs, are selected in the old Mac, by scanning the genome for homology: scnRNA homologous to Mac sequences are degraded and IES-specific scnRNAs are selected.

The scnRNA model opened a lot of questions. First, it was unclear how scnRNA-Twi1p complex are transported in the parental Mac at mid-stages of conjugation. A recent study (Noto et al., in press) has shown that Mac localization of scnRNA-Twi1p complexes is a tightly controlled process, requiring two steps: 1) endonuclease cleavage ("slicing") of scnRNA passenger strand by the Slicer activity of Twi1p and 2) recognition of the mature Twi1p-scnRNA complex - in which scnRNA are single stranded- by a novel Twi1p-associated protein, Giw1p (Gentleman-In-Waiting-1). As a real molecular gentleman, Giw1p waits for scnRNA-Twi1p complexes to be mature and transport them to the parental Mac. Therefore, Giw1p serves as a gate-keeper that allows only those mature complexes to enter Macronuclei, possibly by interacting with some Mac-specific import machinery.

Another unsolved issue in the scnRNA model was how scnRNA interact with genomic sequences first in the parental and then in the developing Mac. Recently, it has been shown (Aronica et al., 2008) that non-coding RNAs (ncRNAs) are transcribed both in parental and newly developed Macs and Twi1p interacts with these transcripts in both nuclei. Moreover, a Twi1p-associated DExH-box RNA helicase, Ema1p is required for efficient Twi1p–ncRNA interaction. This requirement explains the phenotypes of *EMA1* KO strains, including loss of selective down-regulation of scnRNAs homologous to Mac-destined sequences, loss of H3K9 and H3K27 methylation in the developing new Mac, and failure to eliminate DNA. Therefore, both IES recognition by genome scanning and IES elimination are likely to be mediated by small RNA-nascent RNA interactions. In this respect, DNA elimination in *Tetrahymena* is similar to heterochromatin formation in *S. pombe* and paradoxically requires transcription to establish heterochromatin modification. Interestingly, also in *S. pombe* a putative RNA helicase, Hrr1, associates with the argounaute protein Ago1 and is required for its localization to the centromeric repeats and for heterochromatin formation (Motamedi et al., 2004).

Combining these last findings with previous studies, a refined version of the scnRNA model has been proposed which include Giw1p, ncRNAs and Ema1p (Aronica et al., 2008; Noto et al, in press) (Fig 7). In the early stages of conjugation (soon after mating, in the first 4 hr post-mixing), bidirectional non-coding RNA (ncRNA) transcription occurs in the Mic in meiotic prophase. The resulting double stranded Mic ncRNAs (shown as blue wavy lines in Fig. 7) are processed to ~28 nt scnRNAs by Dcl1p. Then, scnRNAs are transferred to the cytoplasm where they form complex with the Argonaute Twi1p. Twi1p cuts by endonuclease cleavage (slicing) the passenger scnRNA strand, which is removed to produce a mature Twi1p-scnRNA complex specifically recognized by Giw1p. Twi1p-scnRNA-Giw1p complex is transferred into the Mac in the mid stages of conjugation (4 hr to ~7 hr post-mixing) and associates with Ema1p and with two additional novel proteins, CnjBp (a Tex domain, GW repeat and CCHC zinc-finger containing protein) and Wag1p (a GW repeat protein) (Bednenko et al., 2009). In parallel, ncRNAs are made from parental Mac chromosomes (red wavy lines) and homologous scnRNAs base-pair with them and are selectively degraded. Ema1p (drawn in purple) likely functions in this selective elimination of scnRNAs, possibly by unwinding the parental Mac ncRNAs to enhance the scnRNA-ncRNAs interaction. scnRNAs forming complex with ncRNAs may be recognized and digested by a specific ribonuclease or simply lose Twi1p and get exposed to non-specific ribonucleases. Subsequently, Ema1p-CnjBp-Wag1p complex transfers IES-specific scnRNA to the developing new Mac in the late stages of conjugation (>7 hr) and recruits a Policomb-like complex containing a histone methyltransferase, Ezl1p (Liu et al., 2007), to induce H3K9/K27me on the chromatin transcribing homologous ncRNAs (green wavy lines). Ema1p could promote H3K9/27 methylation on IESs by enhancing the interaction between scnRNA and their complementary nascent ncRNAs. Then, the chromodomain proteins Pdd1p and Pdd3p binds methylated H3K9/27me promoting the formation of heterochromatin-like structures referred as dumposome (Madireddi et al., 1996). Finally, an unidentified endonuclease (excisase) eventually removes IESs and rejoins the flanking Mac destined sequences.

4. *Tetrahymena thermophila*: a simple single cell with a complex RNAi machinery

RNA-directed DNA elimination is possibly only one of several RNAi pathways in *Tetrahymena.* The RNAi machinery of *Tetrahymena* make up a big family of proteins, including three different Dicer-like enzymes (Dcl1p, Dcr1p and Dcr2p) (Malone et al., 2005; Mochizuki & Gorovsky, 2005), a single RNA-dependent RNA polymerase (Rdr1p) (Lee & Collins, 2006, 2007) and twelve Argonaute Twi proteins, all belonging to the Piwi clade (Cerutti & Casas-Mollano, 2006; Seto et al., 2007; Couvillion et al., 2009). Moreover, the twelve Twi proteins bind a remarkable variety of sRNA deriving from pseudogenes, DNA repeats including telomeres, RNAs with internal regions of secondary structure, and EST-supported mRNAs with convergently transcribed or paralogous genes (Couvillion et al., 2009). Some of these sRNA are novel, since they do not share common features with other sRNA in other organisms. The different RNAi effectors in *Tetrahymena* may be involved in multiple pathways of RNA silencing- most of which have still to be elucidated.

So far, the two best characterized RNA pathways in *Tetrahymena* are those mediated by scnRNA and by an other class of ~23-24 nt small RNA (Lee & Collins, 2006). As described above, scnRNA are ~28 nt long, are specifically produced during conjugation and are involved in heterochromatin formation and DNA elimination. The sequence features of scnRNA is consistent with their proposed origin from dsRNA precursors and their function in RNAi: they exhibit an A:U ratio of nearly 1:1 and, like miRNAs in plants and metazoans and rasiRNAs in *Drosophila melanogaster* (Lau et al., 2001) (Aravin et al., 2003), show a strong 5’ U bias (Lee & Collins, 2006). The second class is composed of ~23-24 nt sRNAs that are constitutively expressed in *Tetrahymena*. In contrast to scnRNA, the 23-24 nt RNAs are proposed to derive from single-stranded precursors (Lee & Collins, 2006). First, the 23-24 nt sRNAs map in an unphased manner at 12 genome loci to the MAC genome and all sRNAs within a cluster are encoded on the same strand. Interestingly, the majority of those clusters are antisense to predicted protein-coding genes lacking EST support and possibly representing pseudogenes or mobile DNA elements. Moreover, ~23–24-nt sRNA present an A:U ratio of about 2:3 and their strand bias cannot be explained basing on the thermodynamic properties of their predicted duplexes, as observed for miRNAs (Khvorova et al. 2003) and plant trans-acting siRNA (Vazquez et al., 2004). Based on these and other evidences (Lee & Collins, 2007; Lee et al., 2009), it has been proposed that the ~23–24-nt sRNA are produced from single-stranded RNA precursors by the physical and functional cooperation of Rdr1p and Dcr2p and that they may be involved in silencing of pseudogenes and mobile elements. The coupling of synthesis and dicing would explain the strand bias observed for the ~23–24-nt sRNA.

Rdr1p forms distinct Dcr2p-associated RDRCs, by exclusive interaction with four tightly Rdr1p-associated proteins: two non-canonical RNA-specific nucleotidyl transferases (rNTrs), Rdn1p or Rdn2p, and two additional subunits, Rdf1p and Rdf2p (Lee et al., 2009). In vitro, the three different RDRCs resolved so far interact physically and functionally with Dcr2p to produce ~23–24-nt sRNA. In vivo, they may modify either ssRNA templates to enhance target recognition or dsRNA products to impact Piwi loading or sRNA turnover. However, the diverse Rdr1p-associated proteins show distinct expression profiles in vegetative and conjugative stages and associate with different specificity with each other. Moreover, their disruption determines different cellular phenotypes and impact differently the overall accumulation of the ~23–24-nt sRNA. Therefore, some Rdr1p-complexes in *Tetrahymena* may be involved in the biogenesis of other small RNAs.

The simple single cell *Tetrahymena* seems therefore to harbor a complex RNAi-machinery with possible roles in multiple pathways. Further studies are required to better characterize the function of the different RNAi players in *Tetrahymena*, their mutual interactions and their associated RNAi pathways.

5. Conservation of RNAi mediated heterochromatin formation

In *Tetrahymena thermophila* small RNA direct heterochromatin formation and DNA elimination. Small RNA pathways involved in heterochromatin formation and chromatin dependent gene silencing (CDGS) are conserved in other eukaryotes. In all the RNA silencing pathways characterized so far Argonaute proteins interact with chromatin or DNA associated molecules. This interaction is often mediated by chromatin adaptor (CAD) proteins containing the conserved glycine and tryptophan (GW)-motif, which bind the PIWI domain of Argonaute proteins and direct them to a variety of different targets (Fig. 5).

In *S. pombe*, the Dicer Dcr1, the Argonaute Ago1 and the RdRP Rdp1 are required for methylation of histone H3 on lysine 9 and silencing at centromeres and mating type loci (Hall et al., 2002; Volpe et al., 2002) (Reinhart & Bartel, 2002). Moreover, in *S.pombe*, RNAi and heterochromatin are physically linked together through RITS which contains centromeric siRNAs, the Argonaute Ago1, the chromodomain protein Chp1, and the glycine and tryptophan (GW)-motif-containing protein Tas3 (Verdel et al., 2004; Buhler et al., 2008; Moazed, 2009). Tas3 is a CAD protein linking Ago1 to Chp1 (Till et al., 2007; Debeauchamp et al., 2008) which binds H3K9me and recruits to chromatin RITS bound to nascent transcripts (Noma et al., 2004). Moreover, RITS recruits a RDRC to non-coding transcripts produced from centromeric repeats (Noma et al., 2004; Sugiyama et al., 2005). *S. pombe* RDRC harbors the RNA-dependent RNA polymerase (RdRP) Rdp1, the putative RNA helicase Hrr1 and the non-canonical polyA polymerase Cid 12 (LaCava et al., 2005; Wyers et al., 2005). Since the physical association of RITS and RDRC depends on siRNA and Clr4, it may occur on chromatin and require histone H3K9 methylation.

In *A. thaliana*, RNA-directed DNA elimination requires the Argonaute AGO4, the Dicer DCL3 and a plant-specific heterochromatin-dependent RNA polymerase, Pol IV (Herr et al., 2005; Li et al., 2006; Pontes et al., 2006). In this organism, RNAi and heterochromatin are physically coupled by the association of AGO4 to Pol IV subunit NRPD1B, which contains a GW motif able to bind AGO4 (El-Shami et al., 2007). Through this binding, AGO4 localizes to nascent non-coding RNAs transcribed by Pol IV and recruits enzymes involved in histone H3K9 and DNA methylation (Moazed, 2009).

In *Drosophila*, heterochromatic silencing requires both components of the RNAi machinery and chromatin proteins. Two major piRNA-producing loci, flamenco and X-TAS (transposable P elements inserted in telomeric-associated sequences on the X chromosome) control transposable elements by a mechanism requiring both PIWI and HP1 (Reiss et al., 2004; Sarot et al., 2004). Moreover, in the *Drosophila* male germline, silencing from the heterochromatic *Suppressor of Stellate* repeats (*Su(ste))* involves 25-27 nt siRNA and depends on both the Argonaute homolog Aubergine (Sting) and the putative RNA helicase Spindle-E (Homeless). Similarly, repeat-induced gene silencing in somatic cells requires several RNAi factors, including PIWI and AGO2, as well as components of the Polycomb group (PcG) of genes (Pal-Bhadra et al., 2004). Finally, the RNAi machinery is also required for heterochromatin formation at *Drosophila* centromeres, association of the heterochromatin protein HP1 and transgene silencing at pericentromeric heterochromatin. In *Drosophila,* PIWI is recruited to chromatin by direct association with HP1, mediated by the HP1 chromo shadow domain (CSD) and PIWI CSD-binding motif (PXVXL, where X is any amino acid) (Brower-Toland et al., 2007). The PIWI–HP1 complex may act as a RITS, in which piRNAs target nascent RNA transcribed from repeated DNA elements and tether them to chromatin. The PIWI–HP1 complex has been so far implicated in heterochromatin assembly and repeat-induced silencing while its involvement in trasposon silencing in the germline remains still unclear (Moazed, 2009).

In *C. elegans*, TGS in somatic cells depends on both RNAi effectors and heterochromatin components including the Dicer *dcr1*, the Argonaute proteins *rde-1* and *rde-4,* the RdRP *rrf-1* as well as HP1 homologs and histone modifiers (Grishok, 2005). In the germline an example of heterochromatic silencing has been described during meiosis (Allis-Jenuwein-Reinberg, 2006): unpaired sequences, such as the X chromosome in males, are silenced by a mechanism which involves H3K9me2, requires a RNA-dependent RNA polymerase (Maine et al., 2005) and resembles meiotic silencing of unpaired DNA (MSUD) in *Neurospora (Shiu et al., 2001).*

Finally, in mammalian cells antisense RNA has been implicated in imprinting and X inactivation (Allis-Jenuwein-Reinberg, 2006) even if a direct role for RNAi in heterochromatin formation in mammalian cells remain to be proved. Nonetheless, introduction of siRNAs into cancer cell lines results in dimethylation of H3K9 on homologous promoters (Ting et al., 2005) and in some cases in DNA methylation (Morris et al., 2004). Moreover, endogenous siRNAs mapping to intergenic regions have been recently identified in *D. melanogaster* and mammalian cells (Ghildiyal et al., 2008; Kawamura et al., 2008) and may be involved in regulation of chromatin structure and function.

All the RNAi pathway described above may have been originally involved in genome defense against extra-chromosomal and/or genome integrated parasitic sequences, by targeting both transcript degradation as well as locus-specific histone modifications (Cerutti & Casas-Mollano, 2006). DNA elimination in *Tetrahymena* may also have originated from the typical RNAi-mediated heterochromatin formation by adding a final step: elimination of the heterochromatic regions. Under selective pressure, some pre-existing enzymatic machinery involved in DNA transposition, recombination or repair may have specialized to provide this final step and delete the silenced sequences (Yao & Chao, 2005).

6. DNA elimination in *Tetrahymena:* evolutional origin and future perspectives

The biology of nuclear dimorphism in *Tetrahymena* may explain what could have been the evolutional advantage of DNA elimination. Each of the two different nuclei in *Tetrahymena* performs a specialized genetic task: the Mac is responsible for most gene expression and the Mac ensures faithful genome transmission. Therefore, consistently with the ancestral function proposed for all the eukaryotic RNAi pathways, the first advantage of DNA elimination may have been to provide a defense mechanism to ensure that invading genetic elements present in the Mic are removed from the somatic genome during conjugation. This hypothesis is consistent with the observation that most IESs are related to transposons and suggests that the deletion process may derive from the transposition mechanism. Therefore, RNAi-directed DNA elimination in *Tetrahymena* may be a genome immune system against parasitic genetic elements invading the germline. Beyond genome defense, an additional advantage of DNA elimination in the Mac may be the reduction of the burden of DNA replication and the simplification of gene expression as a result of heterochromatin elimination (Yao & Chao, 2005).

Another intriguing question is why IESs are retained in the Mic, instead of being eliminated for good. One possible explanation is that the Mic is specialized for genome transmission and IESs, similarly to heterochromatin in other organisms, may play a key role in this process with possible function in mitosis and meiosis. Indeed, some IESs may act as centromers on Mic chromosomes, as suggested by the observation that Mics contain IESs and have centromeres, while Macs lack IESs and show acentric chromosomes (Coyne et al., 1996; Mochizuki & Gorovsky, 2005).

There are many unsolved issues in DNA elimination. These include: regulation of non-coding RNA transcription, biogenesis and nuclear targeting of scnRNA, mechanism of scnRNA selection for the precise targeting of IESs and enzymatic activity involved in IES excision. In this study, I intended to provide new insights in these processes and to help to clarify the evolutional relationships between DNA elimination in *Tetrahymena* and RNA silencing pathways in other organisms.

Results

Identification and characterization of Conjugation-induced (*COI*) genes

To better understand the RNA-directed DNA elimination process in *Tetrahymena thermophila*, I aim to identify novel genes involved in this process. DNA elimination occurs during the sexual reproduction process, conjugation, and most of the genes known to be required for DNA elimination are specifically expressed during conjugation (Madireddi et al., 1996; Nikiforov et al., 2000; Mochizuki et al., 2002; Malone et al., 2005; Mochizuki & Gorovsky, 2005; Aronica et al., 2008; Bednenko et al., 2009; Noto, Submitted). Therefore, I intended first to identify conjugation-specific genes and then determine if those genes could be involved in DNA elimination.

In order to identify conjugation-specific genes, I and other members from my laboratory compared *Tetrahymena* ESTs from vegetative cells (asexually growing cells and starved cells) with those obtained from conjugating cells in silico. We selected the 58 genes most abundantly represented in the conjugative EST and absent in the vegetative ones (Fig. 8). Among these candidate genes, 7 of them (*TWI1*, *PDD1*, *CNJB*, *PDD2*, *PDD3*, *EMA1* and *LIA1*) have already been reported to be involved in DNA elimination (Madireddi et al., 1996; Nikiforov et al., 2000; Mochizuki et al., 2002; Malone et al., 2005; Mochizuki & Gorovsky, 2005; Yao et al., 2007; Aronica et al., 2008; Bednenko et al., 2009;). This indicates that our shortlist of potentially conjugation specific genes is enriched in genes involved in DNA elimination. On the other hand, two of the 58 genes, TTHERM_00459230 and TTHERM_00300660 encode Dmc1 and Mnd1 homologues respectively, which are most probably involved in meiosis (Mochizuki et al., 2008). This suggests that the list could also include genes involved in meiosis and other conjugation specific events. Moreover, the gene TTHERM_00077230 encodes the second largest subunit of RNA polymerase II (Rpb2p), which must have general essential functions in gene expression and thus must be expressed in vegetative stages. This elicits the possibility that some genes in the list are actually not expressed exclusively in conjugation, despite their absence in the vegetative ESTs. Therefore, to find conjugation-specific genes, we first analyzed mRNA expressions of the candidate genes.

We excluded 10 genes listed above (7 known DNA elimination genes, 2 meiosis genes and *RPB2*) from our analysis and studied expression of the other 48 candidate genes by RT-PCR using total RNA extracted from exponentially growing, 16 h starved or conjugating cells (Fig. 8). Twenty-two genes were confirmed to be conjugation-specific, since they were exclusively expressed in conjugation, with no detectable expression in vegetative stages (Fig. 8, 9). These results were consistent with a recently reported mRNA expression analysis using whole genome microarray (Miao et al., 2009). We named these 22 conjugation-specific genes *COI* (COnjugation-Induced) genes.

COI genes can be classified into three classes according to their expression patterns. Class I *COI* genes (*COI1*, *2* and *20*) are mainly detected at early stages (2-6 hr post-mixing) of conjugation, suggesting that they are expressed only from parental Mac. Class II *COI* genes (*COI4, 5, 6, 9, 10, 11, 12, 14, 17* and *18*) are detectable throughout conjugation. This indicates that they are either transcribed from both parental and zygotic macronucleus or their transcripts from parental Mac stably persist until late stages of conjugation. Class III *COI* genes (*COI3*, *7, 13,15,16, 19, 21* and *22*) are expressed only after 6 hr post mixing (Fig. 8 and 9). These genes are probably transcribed from zygotic (newly developed) Mac.

Current status of the functional study of *COI* genes

To understand the roles of *COI* genes, we intended to produce knockout (KO) strains for these genes. There are two ways to produce KO strains in *Tetrahymena*: somatic KO and germline KO. To create somatic KO, one of the ~45 copies of the target gene in the Mac is first disrupted by homologous recombination. Since chromosomes are randomly segregated to the daughter cells during amitotic division of Mac in vegetative growth, strains in which all ~45 copies of the gene are disrupted can be obtained by stepwise selection in increasing concentration of a selection drug. Thus, in somatic KO strains, the target gene is intact in the Mic and in the zygotic Mac derived from it. To create germline KO, one of the two copies of the gene in the Mic is first disrupted by homologous recombination. Then, two heterozygote strains are mated to produce homozygote KO strains. Since the Mic generates both new Mic and new Mac of daughter cells, all copies of the target gene are disrupted in both the micro- and Mac in the germline KO strains.

To study the function of Class III *COI* genes, which are expressed from the zygotic Mac, we need to create germline KO strains. However, producing germline KO is still technically demanding. Therefore, we decided to focus on Class I and II *COI* genes. Class I *COI1* has been shown in a parallel study in our group, to encode for a protein, named Giw1p, which associates with Twi1p and is essential for its Mac localization (Noto et al. in press). Class II Coi4p, Coi5p and Coi11p have been shown to localize at the boundary of two mating cells during conjugation (Kensuke Kataoka, unpublished). These three proteins may play roles in mechanically connecting two cells during conjugation and are probably not involved in DNA elimination process. We did not study these genes further.

Until now, I have succeeded to produce somatic KO strains for *COI6*, *COI9* and *COI12.* In our group, KO strains for two other *COI* genes, *COI2* and *COI8*, have also been produced. Complete replacement of the somatic target loci with the KO constructs was confirmed by Southern hybridization (Fig. 10-13 and 15).

Since most of the genes known to be essential for DNA elimination (*DCL1, EZL1, PDD1, PDD2,* and *TWI1*) are also required for producing viable sexual progeny (Coyne et al., 1999; Nikiforov et al., 2000; Mochizuki et al., 2002; Mochizuki & Gorovsky, 2005; Liu et al., 2007), we first analyzed the progeny viability for each of the *COI* KO strains (results are summarized in Fig. 9). *COI2* and *COI6* KO strains produced viable sexual progeny (Fig. 10C and 11C). *COI9* KO strains showed an unidentified defect in early conjugation and majority of pairs aborted mating during or right after meiosis (Fig. 14B). However, despite this defect, *COI9* KO strains could produce a few viable progeny (Fig. 14A). Therefore *COI2, COI6* and *COI9* are most probably dispensable for DNA elimination and we did not analyze these genes further. It is still possible that these genes have important functions in DNA elimination and other genes act redundantly with them although we could not find any genes homologous to these *COI* genes in *Tetrahymena* Mac genome. *COI8* KO strains can not produce viable sexual progeny (Fig. 12C) and our further analysis confirmed that *COI8* is indeed required for DNA elimination process (Henriette M. Kurth and Kazufumi

Mochizuki, unpublished). Functional analysis of *COI8* will be described elsewhere. In this study, I report the functional analysis of *COI12,* which is required for producing viable progeny and for carrying out DNA elimination (see below).

Parental *COI12* is essential for the generation of viable conjugation progeny

To address whether *COI12* has a function in DNA elimination, somatic *COI12* knockout (KO) *Tetrahymena* strains were constructed (Fig. 15). The entire *COI12* coding sequence was replaced by homologous recombination with a drug resistance cassette, *neo3*, which confers paromomycin resistance in *Tetrahymena* (Fig. 15A). Complete replacement of the Mac *COI12* loci with the KO construct was confirmed by Southern blot (Fig. 15B). Total DNA isolated from wild-type (WT) or *COI12* KO strains was digested with BglII and EcoRV and hybridized using a probe against the 3' end of *COI12* open reading frame. Since a BglII site is present in the KO construct but not in the WT locus, products of different size are generated by the restriction digestion of WT and KO loci, and fragment of different length (5.2 and 1.3 Kb for WT and KO loci, respectively) are recognized by the probe described above. As shown in Fig. 15B, a strong 1.3 kb band corresponding to the KO locus and a faint 5.2 kb band corresponding to the wild-type *COI12* locus, were detected in the *COI12* KO strains. The faint band probably derived from the wild-type *COI12* copies in the Mic, which were not eliminated in the somatic KO. I also studied expression of Coi12p in mating WT and *COI12* KO strains by Western blot using anti-Coi12p antiserum (Fig. 15C). No COI12p was detected in *COI12* KO strains confirming the complete replacement of *COI12* loci with the KO construct.

As for the other KO strains carrying deletions of *COI* genes (See above), I first analyzed progeny viability of *COI12* KO cells during conjugation (Fig. 16A). At 6~8 hr post-mixing, single mating pairs of WT and *COI12* KO strains were placed into drops of SPP medium, incubated for ~60 hr at 30°C, and examined for cell growth. Most of WT cells could resume growth after re-feeding, in contrast most of *COI12* KO cells died. Moreover, the few surviving *COI12* KO cells were actually parental cells that did not complete conjugation, since they did not express the marker specific for the newly developed Macs (6-methyl purine). Therefore, *COI12* KO strains fail to produce viable progeny. This finding elicited the possibility that *COI12* could have a role in DNA elimination.

Then, I further characterized developmental profile of *COI12* KO strains, which is also frequently affected when a gene required for DNA elimination is disrupted. *COI12 KO* strains mate normally (Fig. 16B) but stage M1-M2 and L1-L2 (See Fig. 2 for *Tetrahymena* conjugation stages) are delayed (Fig. 16B), probably due to the defects in scnRNA cytoplasmic function and in heterochromatin formation (see below). Nuclear differentiation of *COI12* KO strains appears cytologically normal and exconjugants containing the new Macs and Mics are produced (even though they are not viable, as showed above). However, the majority of *COI12* KO progeny retains two Mics (Fig. 16B, stage L2) even at 30h post-mixing. In contrast, the progeny of wild-type cells eventually eliminates one of the two Mics (Fig. 16B, stage L3). The arrest at 2Mac2Mic (L2) stage has been reported in strains carrying deletion of genes essential for DNA elimination, such as *TWI1, EMA1,* and *PDD1* KO strains (Coyne et al., 1999; Mochizuki et al., 2002; Aronica et al., 2008), indicating that completion of genome rearrangement is required for completion of conjugation. Therefore, the developmental arrest of *COI12* KO may as well be caused to a failure to carry out DNA elimination.

COI12 is required for DNA elimination

The lack of viable progeny and the late stage developmental arrest in *COI12* KO strains may be due to defects in DNA elimination. To confirm this hypothesis, DNA elimination in the progeny of *COI12* KO strains was analyzed by fluorescence *in situ* hybridization (FISH) using a probe directed against the Tlr1 element (Fig. 17). Tlr1 element is a retrotransposon-like IES that is repeated ~30 times in the Mic genome and is completely removed from the new Mac as a result of DNA elimination (Wuitschick et al., 2002).

I first optimized FISH method using Tlr1 probe to assess DNA elimination event in WT and *TWI1* KO strains (Fig.17 A and B). Since DNA elimination occurs at around 14 h post-mixing in wild-type cells, mating WT and *TWI1* KO strains were fixed at 48 hr post-mixing, in order to exclude the influence of any developmental delay on the observed phenotype. In 99.0% of WT progeny the FISH signal for *Tlr1* loci was detectable only in Mic but not in the newly developed Mac. In contrast, in 99.2% of *TWI1* KO strains, the Tlr1 FISH signal was still present in the new Mac. This is consistent with the fact that *TWI1 is essential for* DNA elimination (Mochizuki et al. 2002). Therefore, the FISH analysis is a valid method to study DNA elimination*,* at least at the Tlr1 loci.

I next analyzed DNA elimination of Tlr1 element in the progeny of wild-type (WT) and *COI12* KO strains by FISH analysis (Fig. 17 C and D). At 48 hr post-mixing, 99.5% of exconjugants (progeny) from wild-type cells showed Tlr1-FISH staining only in the Mic. In contrast, 99.0% of exconjugants from *COI12* KO strains showed Tlr1-FISH staining in both the micronuclei and the newly developed macronuclei, indicating failure to eliminate this IES. It is again important to know that DNA elimination occurs at around 14 hr post-mixing in the wild-type cells. Thus, even though *COI12* KO strains show a slight developmental delay (Fig. 16B), the presence of Tlr1 elements in the new Macs of *COI12* KO strains at 48 h post-mixing is not caused by this developmental delay. Therefore, failure to eliminate Tlr1 loci in *COI12* KO strains is most likely due to direct inhibition of the DNA elimination process.

COI12 is required for heterochromatin formation in new developing macronuclei

It is known that heterochromatin formation involving accumulation of the methylated histone H3 at lysine 9 (H3K9me) and lysine 27 (H3K27me), is prerequisite of DNA elimination (Madireddi et al., 1996; Taverna et al., 2002; Liu et al., 2004; Liu et al., 2007). Since *COI12* is required for DNA elimination, I also analyzed heterochromatin formation in the progeny of *COI12* KO strains by immuno-fluorescence staining using anti-trimethylated H3K9 (H3K9me3) and H3K27 (H3K27me3) antibodies (Fig. 18).

In the progeny of WT cells H3K9me3 and H3K27me3 specifically accumulate at late stage of conjugation (~8 h post-mixing) in developing new macronuclei- exactly when and where DNA elimination takes place. In contrast, no H3K9me3 and H3K27me3 were detected in the progeny of *COI12* KO strains. These results indicate that *COI12* acts upstream of heterochromatin formation in the pathway leading to DNA elimination.

COI12 localizes in both cytoplasm and nuclei

In order to analyze localization of Coi12p, rabbit antibodies against recombinant GST-Coi12p expressed in *E. coli* were generated. The anti-Coi12p antibody recognizes a protein migrating at ~90 kDa position on Western blot made from wild-type cell lysates (Fig. 15, arrow). This protein was not detected from *COI12* KO cells (Fig. 15C), confirming that the protein is Coi12p. The migration of Coi12p in SDS-PAGE was slower than expected from the predicted molecular weight of Coi12p (64 kDa). This slower migration is possibly caused by the very high lysine content (61/546 amino acids) and/or by the acidic nature (estimated pI = 4.81) of Coi12p. A similar migration delay was also observed when recombinant GST-Coi12p was separated by SDS-PAGE and stained with Comassie blue (data not shown).

Then, I analyzed the localization of Coi12p by indirect immuno-fluorescent staining using the anti-Coi12p antibody. WT conjugating cells fixed at 3h, 6h and 9h post-mixing were used for the staining (Fig. 19 A-C). In WT strains, Coi12p localizes in cytoplasm at early stages of conjugation (3 hr). Coi12p is detected mainly in the parental Mac at mid conjugation (6 hr) and in new Macs at late stages (9 hr). The same localization pattern of Coi12p was observed using *COI12*-GFP (Fig. 19, D-E) or *COI12-HA* tagged strains and anti-HA antibodies (Fig 19, F). However, in both *COI12-GFP* and *COI12*-HA strains no fluorescent signal is detected at late stages of conjugation. This can be due to the fact that *COI12-GFP* and *COI12*-HA loci are only in the parental Mac which is destroyed at late conjugation (around 10 hr post-mixing), while *COI12* is likely expressed in both parental and new Macs like other Class II *COI* genes. Interestingly, Coi12p localization pattern mirrors that of scnRNA in complex with the Argonaute protein Twi1p, indicating that Coi12p may promote formation and/or function of these nucleoprotein complexes during conjugation.

COI12 is required for both Mac localization and stable accumulation of Twi1p

Since the Argonaute protein Twi1p plays a pivotal role in DNA elimination (Mochizuki et al., 2002) and *COI12* is required for this process, I asked if the localization of Twi1p was affected in *COI12* KO cells. Localization of Twi1p was analyzed by indirect immunofluorescent staining using an anti-Twi1p antibody. In wild-type cells, Twi1p was detected both in the cytoplasm and the parental Mac at early conjugation stages (Fig. 20A for meiotic prophase) and then localized to the parental Mac at mid-stages of conjugation (Fig. 20C for nuclear exchange stage). In contrast, in *COI12* KO cells Twi1p localized only to the cytoplasm at early conjugation stages (Fig. 20B). This result clearly indicates that *COI12* is essential for the Mac localization of Twi1p.

In mid- (Fig. 20D) and late (data not shown) stages, Twi1p was hardly detected in *COI12* KO strains by immunofluorescent staining. This result was also confirmed by Western blotting (Fig. 20E). Although comparable amount of Twi1p was detected at 2 hr post mixing (meiotic prophase) in both WT and *COI12* KO cells, amount of Twi1p at 4 hr post mixing (meiosis) was significantly lower in *COI12* KO cells than in wild-type cells and Twi1p became undetectable in *COI12* KO cells at later stages. These results indicate that *COI12* is also required for the stable accumulation of Twi1p.

Since Twi1p is essential for heterochromatin formation and DNA elimination (Mochizuki et al. 2002, Liu et al. 2004), the mislocalization and/or the destabilization of Twi1p provide an explanation for the defects in heterochromatin formation and DNA elimination observed in *COI12* KO strains.

***COI12* is required for scnRNA accumulation and loading into Twi1p**

It is known that, in WT cells, scnRNA are produced by the Dicer-like protein Dcl1p in the Mic and then transported into the cytoplasm, where they make complex with the Argonaute protein Twi1p. Then, Twi1p removes one of two strand of scRNA by endonuclease (Slicer) cleavage. Finally, Giw1p (former called Coi1p) binds specifically this mature scnRNA-Twi1p complex and brings it to the parental Mac. Each of these steps is required for the Mac localization of Twi1p, since in *DCL1* KO, Slicer-deficient *TWI1* mutant and *GIW1* KO strains, Twi1p is not detected in the parental Mac. As a similar defect in the Mac localization of Twi1p is observed in *COI12* KO strains, I suspected that *COI12* could have a function in one of the steps leading scnRNA and Twi1p to localize to the parental Mac.

I first asked if Coi12p, like Dcl1p, could be required for scnRNA production. I extracted total RNA from wild- type (WT) and *COI12* KO cells at different time points during conjugation (2, 4, 6, 8, 10 and 12 hr post mixing) and analyzed it by polyacrilamide gel electrophoresis, followed by GelRed staining (Fig. 21). In WT strains, scnRNAs accumulate from early to late stages of conjugation (Fig. 21, left). In contrast, in *COI12* KO cells scnRNAs are expressed at early conjugation stage, but they become undetectable in mid and late conjugation (Fig. 21, right). Thus, Coi12p is not required for production of scnRNAs, but for their stable accumulation.

It is known that Twi1p is also required for the stable accumulation of scnRNAs (Mochizuki et al., 2002). This is probably because scnRNAs are instable and/or exposed to the action of nucleases when they are not associated with Twi1p. Therefore, I then asked if the defect in scnRNA accumulation observed in *COI12* KO strain could be due to a defect in scnRNA loading into Twi1p. Twi1p was immunoprecipitated from WT and *COI12* KO cells at 2.5 hr post-mixing using an anti-Twi1p antibody and co-precipitated RNA was analyzed by denaturing polyacrylamide gel electrophoresis (Fig 22A). It is important to note that, although Twi1p and scnRNAs are destabilized in the absence of *COI12* (Fig. 20E and Fig. 21), comparable amount of Twi1p and scnRNA is expressed in wild-type and *COI12* KO cells at 2 hr post-mixing, when I performed this experiment. Indeed, comparable amount of Twi1p was precipitated from WT and *COI12* KO cells (Fig. 22A, bottom) and comparable amount of scnRNAs was detected in cell lysates from WT and *COI12* KO strains (Fig. 22A, input). Consistently with my hypothesis, no detectable scnRNA was precipitated with Twi1p from *COI12* KO strains, while scnRNAs were clearly co-precipitated with Twi1p from WT cells (Fig. 22A). These results indicate that *COI12* is required for the loading of scnRNA into Twi1p. Since early expression of scnRNA is normal in *COI12* KO strains, Coi12p must act downstream of scnRNA production and upstream of scnRNA loading.

COI12 is dispensable for the export of scnRNAs from the Mic to the cytoplasm

ScnRNAs production is believed to occur by the action of the Dicer-like protein Dcl1p in the Mic (Malone et al., 2005; Mochizuki & Gorovsky, 2005), while scnRNA loading into Twi1p takes place probably in the cytoplasm

(Mochizuki et al., 2002; Noto, Submitted). Therefore, I suspected that the lack of scnRNA loading into Twi1p observed in *COI12* KO cells could be due to a defect in scnRNA export from the Mic to the cytoplasm.

I tested this hypothesis by analyzing amount of nuclear and cytoplasmic scnRNA in WT and *COI12* KO cells (Fig. 22B). WT and *COI12* KO strains were harvested at early conjugation (2.5 hr post-mixing) and nuclear and cytoplasmic fractions were separated. Total RNA was extracted separately from the nuclear and the cytoplasmic fractions and scnRNA was analyzed by Northern blot using a probe directed against Mi9 scnRNA (Aronica et al. 2008). In WT cells, scnRNAs were detected both in the cytoplasmic and in the nuclear fractions. In contrast, in *COI12* KO cells scnRNAs were detected mainly in the cytoplasmic fraction. U3 snRNA (Fig. 22B, bottom), expected to localize in the nucleus, was indeed predominantly detected in the nuclear fraction of both WT and *COI12* KO cells, confirming that the nuclear-cytoplasmic fractionation was successful.

These results indicate that in *COI12* KO strains scnRNAs are exported normally from the Mic to the cytoplasm. However, they accumulate in the cytoplasm and do not reach the nucleus since they are not loaded into Twi1p (Fig. 22A), which in WT strains import them into the Mac. Therefore, Coi12p must act downstream of scnRNA export from the Mic and upstream of, or directly in the process of, the loading of scnRNA into Twi1p.

Coi12p is a ssRNA-binding protein and is probably not a scnRNA loading factor

I next intended to test if Coi12p is involved in the loading of scnRNA into the Argonaute Twi1p. Two properties were expected for scnRNA loading machinery: 1) an ability to bind to Twi1p and 2) an ability to bind to double-stranded scnRNAs. Therefore, I investigated if Coi12p shows these two properties.

First, I analyzed Twi1p-Coi12p interaction in vitro by GST-pull down assay using recombinant GST-Twi1p and Coi12p. GST-Twi1p and 35S-labeled Coi12p recombinant proteins were incubated on a glutathione sepharose resin and eluted complexes were analyzed by SDS-PAGE and autoradiography. No detectable Coi12p was co-immunoprecipitated with Twi1p (data not shown). Next, I analyzed co-immunoprecipitation of Twi1p with Coi12p in vivo, using *Tetrahymena* conjugating cell lysates prepared at 3 hr post-mixing and anti-Coi12p antibodies. Immnunoprecipitated proteins were analyzed by mass spectrometry. No Twi1p was identified in Coi12p immunocomplexes (Appendix 1). Therefore, Tw1p is likely not associated with Coi12p, although I cannot exclude the possibility that Twi1p-Coi12p interaction was disrupted in the experimental condition used in this study.

Then, interaction between Coi12p and RNA was analyzed by electrophoretic mobility shift assay. This method involves mixing together a protein with radioactive-labeled RNA and run the mixture on a native polyacrylamide gel. If a complex is formed, the bound RNA will move slower than the unbound fraction, creating a shift visible on the gel. I analyzed interaction between recombinant GST-Coi12p and 1) double-stranded 29nt scnRNAs, which possess 2 nt 3’-overhangs mimicking Dicer-processed RNA and in which one of two strands was ^{32}P 5’-end labeled (Fig. 23A), 2) ^{32}P 5’-end labeled single-stranded 29 nt RNA (Fig. 23B), 3) ^{32}P internal labeled single-stranded 100 nt RNA (Fig. 23C), and 4) ^{32}P 5’-end labeled single-stranded 29 nt DNA (Fig. 23D). Unexpectedly, I detected a very low binding between GST-Coi12p and the double-stranded 29nt RNA but a strong association of GST-Coi12p to the single-stranded 29 nt RNA and to the single-stranded 100 nt RNA. The dissociation constants (Kd) governing these interactions are ~3nM for 29 nt ssRNA and ~4nM for 100 nt ssRNA (Fig. 24). The low level of bound RNA detected using double-stranded 29nt scnRNAs was likely due to contamination of ssRNA in the reaction mixture. No interaction between GST-Coi12p and the single-stranded 29 nt DNA (Fig. 23D) was detected, indicating that Coi12p selectively binds to RNA. All these results strongly indicate that Coi12p does not bind to double-stranded scnRNAs but is a single-stranded RNA binding protein with little size preference.

No detectable interaction between Coi12p and Twi1p and preferential interaction of Coi12p to single-stranded but not to double-stranded RNAs suggest that Coi12p is not likely a loading factor for scnRNA into Twi1p but has some other role downstream of scnRNA export from the Mic and upstream of the loading of scnRNA into Twi1p.

Coi12p may be part of a novel RNA-dependent RNA polymerase complex

To gain more insights in the function of Coi12p, I sought to identify proteins associated with Coi12p. Coi12p was immunoprecipitated from WT conjugating cells harvested at 3 hr post-mixing, using the anti-Coi12p antiserum (Coi12-IP) or pre-immunoserum (pre-IP). The purified protein complex was separated by SDS-PAGE and visualized by silver staining (Fig. 25). Several proteins were detected that were co-immunoprecipitated with anti-Coi12p serum but not with pre-immune serum, indicating that proteins specifically associated with Coi12p were enriched.

Then, these precipitated proteins were analyzed by mass spectrometry (Appendix 1). 158 proteins were exclusively identified from the precipitate with the anti-Coi12p antiserum. Since the significance of this association is inversely related to the protein molecular weight (MW) and directly proportional to the number of peptides identified, the co-immunoprecipitated proteins were listed in decreasing MW/peptides-ratio order (Appendix 1). Among these proteins, 34 were related to ribosomal RNA synthesis, transcription and splicing. Also, 7 RNA binding proteins were identified. These findings may be an artifact due to the high abundance of ribosomal RNAs and to the high binding affinity of Coi12p to ssRNA (Fig. 24). Alternatively, these binders may have been immunoprecipitated by cross-reaction of the anti-Coi12p serum with some proteins unrelated to Coi12p, since the serum detects many proteins from *COI12* KO on western blot (Fig. 15C).

Two proteins co-immunoprecipitated with Coi12p are related to the components of RNA-dependent RNA polymerase (RdRP) complexes (RDRCs) in *Tetrahymena* (Fig 26A). It has been shown that a *Tetrahymena* RNA-dependent RNA-polymerase, Rdr1p, exclusively interact with two RNA-specific non-canonical nucleotidyl-transferases, Rdn1p and Rdn2p, and two additional factors, Rdf1p and Rdf2p, to form distinct RDRCs. The three distinct RDRCs resolved so far interact physically and functionally with Dcr2p to produce ~23-24 nt RNA from single stranded-RNA precursors (Lee et al., 2009).

The protein TTHERM_01394310 identified in Coi12p-immunoprecipitate, is similar to Rdn1p and Rdn2p and referred as Rdn3p (Fig. 27A and B). Another co-immunoprecipitated protein, TTHERM_01207560, is homologous to Rdf1p and Rdf2p, and named Rdf3p (Fig. 26A, B). The genes encoding the Rdf proteins are located in tandem in the genome with no intervening open reading frames, suggesting that they are the products of recent gene duplications (Lee et al., 2009). The finding that Coi12p coimmunoprecipitates with proteins similar to the known RDRC components may indicate that Coi12p interacts with a novel RDRC that converts Coi12p-associated ssRNAs to dsRNAs.

Interestingly, the Dicer-related protein Dcr1p was also co-immunoprecipitated with Coi12p (Fig. 26B). The role of Dcr1p in conjugation process is not well studied. An intriguing hypothesis is that dsRNAs produced by Coi12p-associated RDRC may be processed to scnRNAs by Dcr1p. Since another Dicer-like protein, Dcl1p, is essential for the production of scnRNAs, Dcl1p-dependent production of scnRNAs from the bi-directional Mic non-coding transcripts must be prerequisite of this hypothetical RDRC-Dcr1p mediated production of “secondary” scnRNAs. The Dcl1p-processed ”primary” scnRNAs might act as primers to convert the Mic non-coding transcripts to dsRNA. It is known that the loading of small RNAs to Argonaute proteins is tightly linked to their processing in many RNAi-related pathways (Chendrimada et al., 2005; Gregory et al., 2005; Chatterjee & Grosshans, 2009; Wang et al., 2009). Unlike scnRNA produced by the Mic-localizing Dcl1p, secondary scnRNAs would be generated in the cytoplasm and possibly directly loaded into Twi1p. This would explain why scnRNAs are not loaded into Twi1p in *COI12* KO strains.

Coi12p associates with a RNA-dependent RNA polymerase activity

The association of Coi12p with ssRNA and with proteins related to RDRC factors was still puzzling. The missing piece was namely the central component of *Tetrahymena* RDRCs, an RNA-dependent RNA-polymerase (RdRP). Only a single RdRP, Rdr1p, has been identified in the almost completely sequenced Mac genome of *Tetrahymena*. However, Rdr1p was not detected by the mass spectrometry analysis of the proteins co-immunoprecipitated with Coi12p. I also investigated, by a reverse pull down approach, if Coi12p could be co-immunoprecipitated with Rdr1p. For this purpose, *RDR1-HA* strains, expressing HA-tagged Rdr1p at their C-terminal, were harvested at 3 hr post-mixing and Rdr1p-HA-containing protein complexes were isolated using an anti-HA antibody. No co-immunoprecipitated Coi12p was detected by western blotting using the anti-Coi12p antibody (data not shown). Therefore, Rdr1p likely does not interact with Coi12p, although I cannot exclude the possibility that Rdr1p-Coi12p interaction was disrupted in the experimental condition used in this study.

It is possible that some RdRP was co-immunoprecipitated with Coi12p in my experiments, but it was not able to be detected by sequence search using the mass spectrometry analysis data. This could happen if the protein coding sequence is misannotated in the genome database or it is coded by Mic-limited elements, for which sequence information is still not available. Also, it is possible that the hypothetical RdRP does not show detectable homology with other known RdRPs.

To test these possibilities, I then decided to further investigate whether Coi12p is associated to a RdRP activity, by in vitro enzymatic assay (Fig. 27). I incubated Coi12p immunocomplexes with α^{32}P-UTP and cold NTPs in the presence (+) or absence (-) of 100 nt ssRNA template (Fig 28A). The reaction was separated in a denaturing polyacrylamide gel and the products were detected by autoradiography. Template-dependent production of radiolabeled products of heterogeneous length was detected (Fig. 27A, lane2). Since all of these products were eliminated by nuclease S1 treatment (Fig. 27A, lanes 3-5), which specifically degrades single-stranded nucleic acids, they are produced by a nucleotidyl transferase which elongates the template ssRNA and not by a RdRP which should produce dsRNA. This nucleotidyl transferase might be Rdn3p.

Next, I performed a similar experiment using the 100 nt ssRNA template annealed with a 28nt RNA primer which base pairs to the 3' most 28 bases of the template (see schematic drawing in Fig. 28B). In addition to the S1 nuclease sensitive products, ~200 nt S1 nuclease resistant RNA was detected (Fig. 27B). This result indicates that a primer-dependent RdRP activity associates with Coi12p. Since it has been reported that Rdr1p has primer-independent RdRP activity (Lee & Collins, 2007), this result supports the hypothesis that Coi12p associates with a novel RDRC. Notably, primer-dependent RdRP activity associated with a nucleotidyl-transferase activity has been reported for Rdp1 in yeast (Sukujama et al, 2006), Qde1p in *Neurospora* (Makeyev, 2002) and recently in human (Maida, 2009).

Coi12p may associate with a RNA ligase activity

The RdRP assay above detected a ~200 nt product which was twice as long as the 100 nt ssRNA template. How can a RdRP activity generate dsRNAs twice long as the ssRNA template? One possibility is that, after being synthesized by a RdRP activity, the 100 bp dsRNA product is ligated by a RNA-ligase which joins the 5'P end of the template with the 3'OH end of the newly synthesized strand. The resulting product should migrate as a 200 nt RNA on a denaturing gel.

To test this possibility, I first dephosphorylated the primer-annealed template using an alkaline phosphatase and then used it for the RdRP enzymatic assay (Fig. 27C). If a ligase indeed catalyze the end-joining of the 100 bp dsRNA product by 5'P-3'OH ligation, the removal of 5' phosphate from the 100 nt RNA should result in a free-ends 100 bp RNA product in the assay. However, even in this experiment, the nuclease S1-resistant dsRNA product migrates as a single ~200 nt band. A polynucleotide kinase might be therefore also associated with Coi12p and phosphorylate the 5' end of dsRNA substrates. The coupling of polynucleotide kinase and RNA-ligase activities has been reported in other systems. The end-healing and end-sealing steps of the phage T4-induced RNA restriction-repair pathway are performed by two separate enzymes, a bifunctional polynucleotide 5'-kinase/3'-phosphatase and an ATP-dependent RNA ligase, Rnl1. A single enzyme, performs these steps in yeast tRNA splicing (Xu et al., 1990; Apostol et al., 1991) and in baculovirus RNA repair pathway (Martins & Shuman, 2004).

Discussion

In this study, I identified a COnjugation Induced (*COI*) gene, *COI12*, which is specifically expressed during conjugation and required for producing viable sexual progeny. I showed that *COI12* is essential for DNA elimination and acts downstream scnRNA export from the Mic to the cytoplasm and upstream scnRNA loading into the Argonaute Twi1p. Moreover, Coi12p is associated to a primer-dependent RNA-dependent RNA polymerase (RdRP) activity and interacts with two proteins related to RdRP factors, Rdn3p and Rdf3p, as well as with the Dicer Dcr1p. These results suggest that Coi12p may be a component of a novel, conjugative RdRP complex (RDRC) with function in DNA elimination. Possibly, Coi12p-associated RDRC produces dsRNA that are processed by Dcr1p into a second class of scnRNA, secondary scnRNA, which is specifically loaded into Twi1p and act as epigenetic messenger in DNA elimination.

Possible functions of Coi12p and their implication in DNA elimination

The most intriguing finding emerging from the functional analysis of the COnjugation Induced gene *COI12* is the property of Coi12p to bind single stranded RNA and its RNA-dependent RNA polymerase (RdRP) activity.

The current model proposed for the molecular mechanism regulating DNA elimination (Mochizuki et al., 2002; Aronica et al., 2008) does not contemplate any ssRNA in the biogenesis of scnRNAs. Instead, all scnRNAs are believed to be processed from long dsRNA produced by non-coding bidirectional transcription of Mic genome. Since Dcl1p, which is required for the production of scnRNA, localizes in the Mic (Mochizuki &

Gorovsky, 2005), scnRNA synthesis has been so far believed to occur exclusively in the Mic. After being exported into the cytoplasm scnRNAs are loaded into the Argonaute protein Twi1p and continue their nuclear journey to the Mac.

However, the finding that scnRNAs are normally exported into the cytoplasm but are not loaded into Twi1p in *COI12* KO strains, suggests that many steps separate these two events and that the biogenesis of scnRNA is much more complex than expected. It is possible to speculate that scnRNA produced in the Mic and scnRNA on Twi1p constitute two different class of scnRNAs, "primary" and "secondary" scnRNA respectively, and that Coi12p-associated RNA dependent RNA polymerase complex (RDRC) may act in the biogenesis of the latter.

RNA polymerases (RdRPs) are involved in RNA silencing pathways in many eukaryotes, including plants, fungi, nematodes and amoebae (Cerutti & Casas-Mollano, 2006; Wassenegger & Krczal, 2006). They amplify siRNA-mediated silencing by using target RNA as template to synthesize more siRNAs. Nevertheless, no evidence for a role of a RdRP in scnRNA biogenesis has been reported.

Tetrahymena Mac genome encodes a single RdRP, Rdr1p, which exclusively interacts with two RNA-specific non-canonical nucleotidyl transferases, Rdn1p and Rdn2p, and two additional factors, Rdf1p and Rdf2p, to form distinct RDRCs. The three distinct RDRCs resolved so far interact physically and functionally with Dcr2p to produce ~23-24 nt RNA from single stranded-RNA precursors (Lee & Collins, 2007).

The following findings suggest that Coi12p may be a component of a novel conjugative RDRC, with function in DNA elimination (Fig. 26C). First, Coi12p associates with ssRNA (Fig. 23) and with two novel RdRP-related factors, Rdn3p and Rdf3p (Fig. 26 A and B). Secondly, Coi12p coprecipitates with a primer-dependent RdRP activity (Fig. 27). Finally, Coi12p interact with the Dicer protein Dcr1p (Fig. 26A). Therefore, Coi12p-associated RdRP may have a similar role to the vegetative RDRC and Dcr2p in the production of ~23-24 nt RNA, and catalyze the synthesis of dsRNA, which are eventually processed to scnRNAs by Dcr1p. Specifically, Coi12p may act in the recruitment of the RDRC to template RNA by binding the single-stranded non-coding RNA and/or the primary scnRNA. This RdRP-mediated secondary production of scnRNAs could explain why total level of scnRNA does not drop down (Fig. 22) even if Mac-specific scnRNA, which is expected to consist ~85% of total scnRNA initially produced in Mic, are negatively selected during conjugation.

Since Dcl1p is essential for the production of scnRNAs (Mochizuki & Gorovsky, 2005), and normal levels of scnRNAs are accumulated in early stages of conjugation in *COI12* KO cells (Fig. 22), I speculate that Dcl1p-processed “primary” scnRNAs are required for RDRC-Dcr1p mediated production of “secondary” scnRNAs. The primary scnRNAs may act as primers to convert the Mic non-coding transcripts to dsRNAs. If this is the case, the production of “primary” scnRNAs is expected to occur in Mic, where Dcl1p is exclusively localized, while “secondary” scnRNAs might be processed in the cytoplasm where Coi12p is localized (Fig. 19). It is known that the processing of siRNAs and their loading into Argonaute proteins are linked in many RNAi-related pathways (Chendrimada et al., 2005; Gregory et al., 2005; Chatterjee & Grosshans, 2009; Wang et al., 2009). Since Twi1p does not localize to the Mic but first appear in the cytoplasm (Mochizuki et al., 2002), I further speculate that the production of “secondary” scnRNAs is coupled with their loading to Twi1p in the cytoplasm. This could explain why scnRNAs are not loaded into Twi1p in

COI12 KO cells despite the normal production of scnRNAs in their early conjugation stages.

In *Tetrahymena* production of secondary scnRNA may be required to amplify scnRNA before they undergo the selection against Mac-specific sequences, in order to ensure that sufficient amount of IES-specific scnRNA persist until late conjugation and induce DNA elimination. Alternatively, the amplification might be biased for repeated sequences, that are mostly Mic-specific, and thus possibly act as a second mechanism of scnRNA selection.

Possible biosynthetic mechanism of Coi12p associated RdRP

Coi12p-associated RdRP activity produces dsRNA by a mechanism of action involving a primer requirement. Two main mechanisms of dsRNA production have been proposed for RdRPs: *de novo* and primer-dependent synthesis. In the *de novo* mechanism, RdRP generates dsRNA without need of a primer, utilizing as template aberrant RNAs lacking normal processing signals such as a 5'cap or a polyA tail. In the primer-dependent mechanism, RdRP needs a primer to synthesize the complementary strand of target ssRNA. This primer may be represented by a siRNA directed against the target mRNA (Sijen et al., 2001; Baulcombe, 2005; Wassenegger & Krczal, 2006) or a 3' fold-back structure of the template itself- a mechanism known as "back priming" (Laurila et al., 2002; Maida et al., 2009) (Makeyev & Bamford, 2002).

The cellular RdRPs described until now do not show a primer requirement, while several viral RdRPs use both primer-dependent and primer-independent mechanisms (Behrens et al., 1996; Cornell & Semler, 2002). Primer-dependent RdRP activity has been associated to a particular protein architecture, the “right-hand” structure which distinguishes viral from cellular RdRPs and DdRPs (Bressanelli et al., 1999; Butcher et al., 2001). This protein architecture resembles the shape of a "right hand" in which the "thumb," "fingers," and "palm" form a pocket along which the RNA template can move. Interestingly, human TERT, which has recently shown to act as primer-dependent RdRP in human RNAi silencing (Maida et al., 2009), is a right-handed polymerase evolutionarily related to both reverse transcriptases and viral RdRPs (Gillis et al., 2008). Since Coi12p-associated RdRP activity is also primer-dependent, the responsible enzyme may as well have a right-handed protein structure. Identification and characterization of this activity may contribute to confirm the validity of the paradigm proposed for correlating the structure of RdRPs and their mechanism of dsRNA synthesis.

A second unusual feature of Coi12p-associated RdRP activity is the possible involvement of ligation step acting on the dsRNA product. Indeed, by RdRP activity assay I detected a ~200 nt product which was twice as long as the 100 nt ssRNA template. This finding suggests that a 100 bp dsRNA product is first synthesized by a RdRP activity and then ligated by a RNA-ligase which joins the 5’P end of the template with the 3’OH end of the newly synthesized strand. The resulting product should migrate as a 200 nt RNA on a denaturing gel.

Interestingly, the model proposed for secondary RNA production in Drosophila also involves both a RdRP and a RNA ligase step (Lipardi et al 2001). According to this model, multiple siRNA primers align along a ssRNA template strand, are extended by RdRP and eventually ligated by a putative RNA ligase. This mechanism would ensure the generation of dsRNA of sufficient length to be cleaved by RNase III-type activity, which requires a substrate minimum 39 base pairs in length (Elbashiret al., 2001b). Similarly, it is possible that also in Tetrahymena, the ligation step involves in vivo two adjacent strands elongated by RdRP on the same ssRNA template. Alternatively, dsRNA products may be also in vivo end-joined intramolecularly, like it has been observed in vitro. The resulting product would be a hairpin molecule with enhanced stability and/or recognition by processing enzymes, like Dcr1p, or binding proteins, like Twi1p.

Coi12p-associated ligase activity acts also on dephosphorylated templates (Fig. 28C). This finding may indicate that a polynucleotide kinase (Pnk) is as well associated with Coi12p immunocomplex and phosphorylate the 5’ end of dsRNA substrates. However, primary scnRNAs and non-coding transcripts, proposed to act in vivo as ssRNA templates, should be already phosphorylated at their 5’ end. Why is a Pnk involved in dsRNA production by RDRC? Possibly, the Pnk activity associated in vitro with Coi12p may in vivo counteract the action of endogenous phosphatases on 5’ termini of ssRNA templates. Alternatively, Coi12p-associated Pnk activity may catalyze in vivo the inverse reaction and act as a RNA phosphatase. Indeed, in presence of a 5’phosphorylated target and ADP, Pnk can also transfer the phosphate from the nucleic acid, producing a dephosphorylated target and ATP. It is indeed known that dsRNA product of RdRP activity tipically bear a 5’-triphosphorylated end. This terminal modification has been reported for secondary scnRNA in *C.elegans* (Pak & Fire, 2007), and suggested also for dsRNA produced by Rdr1p in *Tetrahymena* (Lee & Collins, 2007). In *C. elegans*, the RNA phosphatase PIR1 associates with DCR1 and likely dephosphorylates the γ- and β-phosphate from the 5’-triphosphorylated termini of secondary siRNA, before their incorporation into RISC (Duchaine 2006). Secondary scnRNA may also be 5’-triphosphorylated and a RNA phosphatase may be required to dephosphorylate them before loading into Twi1p. However, so far no phosphatase has been shown to be essential for DNA elimination. The possibility that such an activity indeed associates with Coi12p needs to be further explored.

Conservation of RNA-dependent RNA polymerase proteins and activities among eukaryotes

The existence of secondary small RNA synthesized by a RdRP has been reported in other eukaryotes. In *C.elegans*, RRF-1 converts RNA targets of primary siRNA into secondary dsRNA bearing a 5'-triphosphate (Pak & Fire, 2007; Sijen et al., 2007), (Axtell et al., 2007; Ruby et al., 2007). In *S. pombe*, RDRC produces secondary siRNA from centromeric non-coding transcripts (Noma et al., 2004; Sugiyama et al., 2005). *Arabidopsis thaliana* RDR6 synthesizes endogenous siRNAs from transcripts recognized by miRNAs (Allen et al. 2005; Yoshikawa et al. 2005, Bu¨hler and Moazed 2007). Even though RdRP genes are absent in the fly *Drosophila melanogaster* and in human genomes, recent evidence suggests that secondary RNA may have function also in these organisms. In *Drosophila,* endogenous siRNA have been described which show phasing and strand bias, like secondary siRNA in *C. elegans*, and a novel RdRP has been identified (Lipardi 2009). Moreover, human telomeric retro-transcriptase (TERT) acts as RdRP producing endogenous siRNA, which are loaded into RISC (Maida 2009).

RdRPs is less broadly distributed among eukaryotes than Ago-Piwi and Dicer-like proteins (Fig. 28). Although RNA-dependent RNA polymerases (RdRPs) are required for RNA silencing in *Arabidopsis thaliana, Neurospora crassa, S. pombe* and *C. elegans* (Baulcombe, 2007), no RdRP homolog has been identified in flies or higher metazoans including humans (Fig. 29). However, many aspects of PTGS in these organisms seem to implicate the existence of some type of RdRP activity also in flies and humans. For instance, aberrant ssRNA are as well a target of RNA-silencing, indicating that they are probably template for dsRNA synthesis in *Drosophila* (Lipardi et al., 2005). Moreover, mathematical models of RNAi (Bergstrom et al., 2003) also suggest that an unidirectional biosynthetic mechanism, such as a primed amplification step involving siRNAs and the target RNA, is required to avoid self-directed reactions that would occur in the catalytic Ago2 based mechanism proposed for RNAi in higher metazoans (Bichler, 2009).

Indeed, Lipardi and Paterson have recently reported that the *Drosophila* RNA polymerase II core elongator complex subunit, Elp1, has both primer-dependent and independent RdRP activity in vitro, interacts with Dicer 2 and is involved in transposon silencing. Elp1-related proteins are conserved in all eukaryotes and recombinant Elp1 from fission yeast, *C. elegans* and human have RdRP activity in vitro. Moreover, another report in human cells has demonstrated that the human telomerase retrotranscriptase (hTERT) is the catalytic subunit of a ribonucleoprotein complex with a back-priming, primer-dependent RdRP activity in vitro. Dicer eventually processes dsRNA produced by TERT complex into siRNA, which are eventually loaded into RISC. Thus, both findings in fly and human suggest that RdRP activity is present also in organism without the canonical RdRP gene.

In this study, I detected a RdRP activity co-precipitated with Coi12p. Although *Tetrahymena* genome harbors the canonical RdRP gene *RDR1*, which is known to be involved in the synthesis of vegetative ~24 nt RNA. I believe Rdr1p is not responsible for the RdRP activity co-precipitated with

Coi12p (Fig. 27) basing on the following two observations. First, I could not detect any interaction between Coi12p and Rdr1p by immunoprecipitation of either Coi12p (Appendix 1) or Rdr1p-HA (data not shown) from conjugative *Tetrahymena* cells. Secondly, although it has been reported that Rdr1p possesses a primer independent RdRP activity, the RdRP activity co-precipitated with Coi12p is primer-dependent (Fig. 27). Possibly, the enzyme responsible for the Coi12p-associated RdRP activity is hard to be detected by sequence search using mass spectrometry analysis data. This may indeed occur either if the protein coding sequence is misannotated in the genome database or if it is limited to the Mic genome, for which sequence information is still not available. Alternatively, Coi12p-associated RdRP activity is harbored by an enzyme other than a canonical RdRP, like in the case of *Drosophila* RNA PolII core elongator factor Elp1. Interestingly, a RNA PolII factor, Rpb8, was identified among Coi12p-associated proteins (TTHERM_00549610 in Appendix1). Even though this finding may be due to a simple cross-reaction of anti-Coi12 antibodies with Rpb8, it may also indicate a true interaction of PolII complex with the RNAi machinery in *Tetrahymena*, possibly mediated by Coi12p and leading to secondary scnRNA production.

The functional analysis of *COI12* has uncovered a new RNAi pathway in DNA elimination involving production of dsRNA by an RdRP activity. Further investigations are required in order to better understand the mechanism and function of this process- including identification of template molecules and effector proteins, characterization of dsRNA products and their interaction with other RNAi players, significance of a RdRP-directed step in DNA elimination. Potential interconnections with RNA silencing pathways in other organisms raise new intriguing issues in the ever-enlarging world of small RNAs.

Current issues and future perspectives in the functional analysis of *COI* genes

The sexual process of conjugation in *Tetrahymena* is characterized by dramatic developmental changes including cell pairing, meiosis, fertilization, chromosome breakage, telomere addition, rDNA amplification and DNA elimination. Many genes required for these processes are specifically expressed or highly upregulated during conjugation (Miao et al., 2009). Among the conjugation-specific genes, most of the genes known to be essential for DNA elimination are represented (Madireddi et al., 1996; Nikiforov et al., 2000; Mochizuki et al., 2002; Malone et al., 2005; Mochizuki & Gorovsky, 2005; Aronica et al., 2008; Bednenko et al., 2009; Noto, Submitted). Turning around this relation, in this study we have identified a pool of COnjugation Induced (*COI*) genes as candidates with a potential role in DNA elimination.

This approach is methodologically simple and involves in silico comparison of conjugative and vegetative EST library to select potentially conjugation-specific genes and confirmation of their expression pattern by RT-PCR to identify true *COI* genes for functional analysis. By using this strategy we have indeed been able to shortlist genes involved in DNA elimination. First, among 58 potentially conjugation-specific genes, 7 genes (*TWI1, PDD1, CNJB, PDD2, PDD3, EMA1* and *LIA1)* have already been reported to be required for DNA elimination (Nikiforov, 2000; Mochizuki, 2002; Mochizuki, 2005; Malone, 2005; Madireddi, 1996; Bednenko, 2009; Aronica, 2008; Yao, 2007; Noto, In press). Moreover, until now we have carried out a functional analysis of 5 *COI* genes (*COI2, COI6, COI8, COI9,* and *COI12*) and 2 of them, *COI8* and *COI12*, have been shown to be involved in DNA elimination. Interestingly, some of the short listed candidate genes have been demonstrated to physically associate with each other (Twi1p with Ema1p, Twi1p with Coi1p/Giw1p, Coi12 with Coi4p). Therefore, the functional analysis of *COI* genes is a valid strategy to identify new players in DNA elimination and uncover their functional interactions, bringing together the small pieces of this puzzling phenomenon.

However, many conjugation-specific genes are not involved in DNA elimination but in other cellular processes occuring contemporaneously during conjugation. For instance, meiosis, fertilization and gene activation during Mac development overlap respectively with the early, middle and later stages of RNAi-mediated IES targeting (Miao et al., 2009). Therefore, some of the identified conjugation specific genes were not further characterized. This is the case of the meiosis genes *DMC1* and *MND1* and of *COI15*, which localizes at the cell boundary and is therefore likely involved in cell pairing (data not shown).

On the other hand, some non conjugation-specific genes, such as *CNJB* and *LIA6*, have been shown to be involved in DNA elimination (Yao et al., 2007; Bednenko et al., 2009). This indicates that these genes have either additional functions or that the correspondent proteins are required in large amounts during conjugation and therefore preparation for their expression begins earlier. Thus, some genes required for DNA elimination may have been methodologically excluded from our analysis. However, since most of the genes required for DNA elimination are conjugation specific, the study of *COI* genes is still a straightforward approach to identify new players involved in this process.

Until now we have not analyzed Class III *COI* genes (*COI3, 7, 13,15,16, 19, 21* and *22*). These genes are expressed only at late stages of conjugation and are probably derived from the zygotic (newly developed)

Mac. Their analysis would therefore require the generation of germline KO, which is still technically demanding. However, since class III *COI* genes are expressed at late stages of conjugation, they are likely enriched in genes required for the late steps of DNA elimination involving the enzymatic machinery responsible for IES excision and MDS ligation. Due to the technical difficulties mentioned above, these late steps have been so far poorly characterized and only a single conjugation-specific gene (*LIA1*) functioning after IES targeting has been identified (Rexer & Chalker, 2007). In contrast, many genes with function in early and mid stages of DNA eliminations (*DCL1, TWI1, EMA1, COI8, COI12, EZL1*) have been characterized, providing important insight into the processes of scnRNA production, selection and targeting of IES. Therefore, the functional analysis of class III genes remains an important issue in the study of *COI* genes. Experimental efforts are currently ongoing to address this issue by simplifying the generation of germline knockouts. This will allow a better understanding of the final stages of DNA elimination, including the mechanisms of IES deletion and re-joining.

Material and methods

Strains and culture conditions

Wild-type B2086 and CU428 strains of *T. thermophila* were provided by Dr. P. J. Bruns (Cornell University, Ithaca, NY). *TWI1* KO strains were described previously (Mochizuki et al. 2002; Mochizuki and Gorovsky 2004b). Cells were grown in SPP medium (Gorovsky et al. 1975) containing 1% or 2% proteose peptone at 30°C. For conjugation, growing cells (5 × 105 cells per milliliter) of two different mating types were washed, prestarved (12–24 h) and mixed in 10 mM Tris (pH 7.5) at 30°C.

Construction of *COI12* KO strains

To make the *COI12* disruption construct, first 5' (369,981 b – 372,433 b of Genbank CH445669) and 3'(209,604b – 210,117b) flanking sequences of *COI12* gene were amplified by PCR with primer sets *COI12*-KO-5'FW (GTGGATCCTAGAAGTTGTGTTGTAAGT)/*COI12*-KO-5'RV (TTCCAGGTCGACAGATCTG ATAAGAGTAACTAGTTTGGTGC) and *COI12*-KO-3'FW2 (AGATCTGTCGACGACTTTCGAGAAAACTCATATCAAGCTTCAG)/ *COI12*-KO-3'RV (GGCTCGAGTGTATGTA TGAGTGTCTCTG).

These 5' and 3' PCR products had a short overlapping sequence at their 3'end and 5'end, respectively. These PCR fragments were then combined by overlapping PCR with *COI12*-KO-5'FW and *COI12*-KO-3'RV. BglII and SalI sites were made in the overlapping region between the two flanking sequences. The combined PCR product was inserted into pBlueScript SK(+) vector and BamHI and XhoI digested *neo4* cassette (Mochizuki, 2008) was then inserted into the BglII and SalI sites. The construct was digested out from the vector backbone by XhoI and SpeI digestion and introduced into the Macs of B2086 and CU428 cells by biolistic transformation as described (Cassidy-Hanley et al., 1997) .The endogenous wild-type Mac *COI12* loci were completely replaced with the disrupted loci by phenotypic assortment (Karrer, 2000) after selection in increasing concentrations of paromomycin (pm). Complete replacement was confirmed by Southern hybridization (see below). For experiments, (B)-3-a-κ and (C)-2-b *COI12* KO strains were used.

Southern hybridization

Total DNA isolated from ~1 x 10^4 growing wild-type (WT) or *COI12* KO strains was digested BglII and EcoRV. The digested DNA was separated on a 0.8% agarose gel, transferred to Hybond N+ membranes (GE Healthcare) by capillary blotting and fixed by UV (1200 J/cm2). Membranes were probed with ~50ng 5'-^{32}P-labeled DNA oligo corresponding to the 5' region of *COI12* (amplified using *COI12*-KO-5'FW/*COI12*-KO-5'RV) in Ultra-Hyb solution (Ambion) for 16 hr at 42°C. Membranes were washed twice in 1 x SSC containing 0.1% SDS for 20 min at 42-60°C and once in 0.5 x SSC and 0.25 x SSC containing 0.1% SDS for 20 min at 60°C. The membranes were finally subjected to autoradiography.

Progeny viability

Viability of progeny was analyzed as described (Mochizuki et al. 2002). Wild-type and *COI6, COI8, COI9, COI12* cells of different mating types (Fig. 11, 12, 14, 16) were mated, individual pairs were placed into SPP drops 5–8 hr post-mixing and cultured at 30°C. Two hours later, drops were examined for swimming cells to exclude from the analysis any cells having been killed during pair isolation. At 60 hr after cloning, drops were examined for growth of cells. The phenotypes of growing cells (>50 cells/drop) were then tested to determine whether the cells had completed or aborted conjugation. To check the completion of conjugation after mating, cells were incubated with 200 µg/ml pm and 1 µg/ml CdCl2 in SPP and scored for sensitivity to pm. This assay depends on the fact that, in all cases examined, one or both of the cells of a conjugating pair contained genes in its macronucleus that conferred pm resistance, but both members of the pair contained pm-sensitive alleles in their micronuclei. Thus, pm-sensitive cells can be obtained only if conjugation is successful. To check for completion of conjugation in wild-type cells (B2086 × CU428), cells were incubated with 15 µg/ml 6-methylpurine (6-mp, Sigma) in SPP and scored for cells resistant to 6-mp. This assay depends on the fact that CU428 cells contain a gene in their micronuclei that confers 6-mp resistance, but both strains contain 6-mp-sensitive alleles in their macronuclei.

Production of anti-Coi12p antibody

We raised antibodies against recombinant Coi12p. To produce recombinant Coi12p, DNA encoding Coi12p that had been optimized for codon usage in *E. coli* was synthesized (GenScript Corporation, Piscataway, USA). Synthesized *COI12* DNA was cloned into a pGEX4T-1 expression vector (Amersham Biosciences) to allow the production of GST-Coi12p fusion protein. GST-Coi12p fusion protein was expressed in *E. coli* strain BL21 (DE3). After cultivation at 37°C to an A600 of ~0.8, cells were incubated with 0.5 mM IPTG for 3 hours at 18°C. Cells were lysed in 500 mM NaCl, 80 mM Tris-HCl (pH 8.0), 0.5% Triton X-100, 0.2 mM PMSF and 1x complete proteinase inhibitor cocktail (Roche). The cell lysate was incubated with glutathione sepharose 4B (Amersham Biosciences) at 4°C. After washing with 500 mM NaCl, 80 mM Tris-HCl (pH 8.0) and 0.1% Triton X-100, the GST fusion proteins were eluted in 160 mM reduced glutathione, 500 mM NaCl and 80 mM Tris-HCl (pH 8.0), and subsequently the buffer was replaced with PBS by dialysis. Recombinant Coi12p was used to immunize rabbits (Gramsch Laboratory) and raise rabbit anti-Coi12p antibodies. For experiment, anti-Coi12p serum #2843 was used.

Western blotting

To prepare whole-cell lysate, 5 x 10^5 starved or conjugative cells at the indicated time points were treated with 10% TCA for 10 min on ice and precipitated cells were lysed in 100 µl SDS-PAGE buffer. Samples were boiled for 5 min and 5-10 µl were resolved by 8% SDS-PAGE. The proteins were transferred to a nylon membrane (Immobilon) by semidry-transfer. The membrane was then blocked with 1% BSA, 0.1% Tween20, 1% skimmed milk, 1 x PBS.

To detect Coi12p, Twi1p, Pdd1p the membranes were incubated with 1/1000 diluted rabbit anti-Coi12p (see above), rabbit anti-Twi1p (Aronica et al. 2008) or rabbit anti-Pdd1p (Millipore) antibodies, respectively, for over-night at 4 °C. Membranes were then washed in PBS supplemented with 0.1% Tween20 (PBST), incubated with HRP-conjugated goat anti-rabbit IgG antibodies at 1/10,000 dilution for 1 hr at room temperature and washed in PBST. The bands were visualized using ECL™ Western Blotting Detection Reagents (GE Healthcare) and exposure to Amersham Hyperfilm ECL (GE Healthcare)

Indirect immunofluorescent staining

Cells (~6 x 10^5) were fixed at different time points during conjugation by the addition of formaldehyde and Triton X-100 (final concentrations of 4% and of 0.5%, respectively). After gentle mixing, the cells were left for 30 minutes at room temperature. Cells were collected by centrifugation (1,600 rpm for 1 min) and resuspended in 500µl of 4% formaldehyde + 3.4%

sucrose. 30 µl of this mixture were spread on a clean SuperFrost Ultra Plus® slide (Menzel) air-dried and either used for immunostaining right away or stored for at – 20 °C).

For immunostaining, slides were incubated first in 1x phosphate buffered saline (PBS) + 0.1% Tween20 (PBT), three times for 10 minutes each, and then in blocking solution (3 % BSA, 10% Normal Goat Serum, 0.1 % Tween 20 in PBS), at room temperature for 2 hr. Primary antibodies [1:800 rabbit anti-Coi12 (see above), 1:500 rabbit anti-Twi1p (Aronica et al, 2008), 1:300 rabbit anti-H3K9me3 (gift from Dr. Thomas Jenuwein, MPI, Germany) or 1:300 rabbit anti-H3K27me3 (gift from Dr. Thomas Jenuwein) antibodies] in the blocking solution were applied under a coverslip overnight at 4 °C. The coverslip was removed and the preparations were rinsed with 1x PBT. 1:1000 diluted Alexa488-conjugated goat anti-rabbit IgG (Invitrogen) in the blocking solution was applied under a coverslip and incubated for ~1 hr at room temperature. The slides were washed three times for 10 minutes in 1x PBT and one time in PBT supplemented with 1 µg/ml DAPI (4', 6-diamidino-2-phenylindole) as a DNA-specific counterstain. Finally, the slides were mounted under a coverslip with 20 µl of ProLong® Gold anti-fade reagent (Invitrogen).

DNA elimination assay by Fluorescent In Situ Hybridization (FISH)

FISH probes for Tlr1-element were produced by nick translation using a protocol adapted from Beatty et al. 2002 and from Jane Bayani and Jeremy A. Squire (Current protocol in Biology Unit 22.4). Labeling reaction were performed by mixing 2 µg pMBR 4C1, pMBR 2 and Tlr IntB plasmid DNA (Wuitschick et al. 2002) with 10X nick translation buffer (0.1 µg/µl BSA, 0.1 M 2-mercaptoethanol, 0.5 M Tris□Cl, 50 mM $MgCl_2$), 50 µM dATP/dCTP/dGTP mixture, 4 µl of Cy3-UTP (GE Healtcare), 10X DNase I solution (150 mM DNAseI, Tris□Cl pH 7.0 50 mM, MgCl2 5 mM, 2-mercaptoethanol 1 mM, BSA 4 mM), 2.5 µl *E. coli* DNA polymerase (New England Biolabs) in 100 µl. Reactions were incubated 120 min at 15 ºC and then placed on ice. Production of labeling products of desired size of 300-700 nucleotides was checked in 2% agarose gel. Once the proper labeling size was achieved, 30 mM EDTA was added to the reaction to stop the action of the enzymes. Unlabeled salmon sperm ssDNA (125 ng/µl) was added to the reaction to suppress unspecific binding of the probe. This DNA mixture was ethanol precipitated at -20 ºC over night and resuspended in hybridization buffer (50% formamide, 10% dextran sulphate, 2x SSC) to a final concentration of 10 ng/µl. Labeled probe were stored at -20ºC up to use for FISH experiment.

FISH was performed using WT, *COI12* KO *[*(B)-3-a-κ and (C)-2-b] and *TWI1* KO cells fixed at 48 hr post-mixing by formaldehyde fixation (as described above for immunostaining) and spread on slides. Preparations were hydrated in H2O followed by treatment with 1 M sodium thiocyanate at 90°C for 15 minutes. Denaturation was performed in 70% formamide, 2x SSC, pH 7.1 for 2 minutes at 68°C. The slides were then rinsed in H2O and air-dried. Labeled Tlr1 probe in hybridization buffer (8 µl per slide) were denatured at 95°C for 3 minutes and cooled on ice. The denatured probe was applied on the slide, covered with cover glass and sealed with rubber cement (Marabu, Hamm, Germany). Slides were denatured again on a hit-block at 80 ºC for 10 min and hybridized for 48 hours at 37°C in a moist incubator. After hybridization, the cover glasses were removed and the slides were washed three times in 50% formamide in 2XSSC (37°) for 2min, 2xSSC (37°C) for 5min and 1xSSC (room temperature) for 5min. Slides were then incubated with 1xPBS, 0.05%Triton, 10ng/ml DAPI for 10 min and finally mounted with ~20µl of ProLong Gold antifade solution (Invitrogen) under a cover slip.

Northern hybridization

Total RNA extracted using TRIzol reagent (Invitrogen) from ~5 x 10^5 starved or mating WT or *COI12* KO cells was separated on 15% denaturing polyacrylamide gels, transferred to Hybond N+ membranes (GE Healthcare) by capillary blotting, fixed by UV (1200 J/cm2), and baked for 2 hr at 80°C. Membranes were probed with 10 pmol random primed ^{32}P-labeled DNA oligos corresponding to *TWI1* or *PDD1* in Ultra-Hyb solution (Ambion) for ~24 h at 42°C. Membranes were washed twice in 2 x SSC containing 0.1% SDS for 20 min at 42°C and once in 1 x SSC and 0.1 x SSC containing 0.1% SDS for 20 min at 65°C. The membranes were finally subjected to autoradiography.

Analysis of nuclear and cytoplasmic scnRNA

WT and *COI12* KO [(B)-3-a-κ and (C)-2-b] mating cells (~1 x 10^9) were harvested at 2.5 hr post-mixing and resuspended in 10.5 ml of lysis buffer including 10mM TRIS-HCL pH 7.5, 5mM $MgCl_2$, 10mM KCl, 0.34mM sucrose and 0.7% octanol. Lysate was prepared by stirring the cell suspendion for 10 min on ice and then first mixed with 12.5 ml of buffer A [18mM TRIS-HCL pH 7.5, 86mM KCl, 23mM NaCl, 3mM EDTA, 0.75 mM EGTA, 2.14M sucrose, 0.23mM spermin-HCl (Mr 348.18), 0.16 mM Spermidine (276ul/l = 1mM), 23 mM BME] and finally layered on 3 ml of buffer B [15 mM TRIS-HCL pH 7.5, 60 mM KCl, 15 mM NaCl, 2 mM EDTA, 0.5 mM EGTA, 2.1 M sucrose, 0.23mM spermin-HCl (Mr 348.18), 0.16 mM Spermidine (276ul/l = 1mM), 15 mM BME] and centrifuged at 25000 rpm and 4 C˚ for 2 hr using AH629 Sorvall rotor. The supernatant and pellet were used as a source of cytoplasmic and nuclear RNA, respectively. The pellet was resuspended in 1 mL of TRIzol

(Invitrogen). Nucleic acids were extracted with phenol/chloroform and precipitated with iso-propanol. Extracted nucleic acids were resuspended in formamide and analysed on a 15% PAGE gel, followed by staining with GelRed (Roche) or Northern blotting, using Mi-9 scnRNA (Aronica et al., 2008) or U3 snRNA as probe. For Northern blot, Mi-9 scnRNA and snU3 RNA were radiolabeled in 10 pmol quanitities with 25 µCi [γ-^{32}P] ATP (Perkin Elmer Life Sciences) and 5 U of Optikinase (USB), according to the manufacturer's directions. Hybridized membranes were exposed 1-3 overnight to a storage phosphor screen (GE Healthcare). Images were captured using a Typhoon Imager (GE Healthcare) and analyzed using ImageQuant software.

Production of recombinant proteins

Recombinant GST and GST-Twi1p were prepared as described (Noto et al. in press). Recombinant GST-Coi12p fusion protein was produced as described above for antibody production. Radio labeled Coi12p recombinant protein was synthesized using PCR-amplified *COI12* synthetic DNA, 35S-methionine and PURExpress In Vitro Protein Synthesis Kit (New England Biolab).

GST-pull down assay

GST, GST-Twi1p (~1 µg) were expressed in E. coli and purified as described above and incubated with 20 µl glutathione sepharose 4B resin (GE Healthcare) in GST pull-down buffer (GPB) (20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 2 mM MgCl2, 2 mM CaCl2, 0.1% BSA) for 30 min at 4°C. The beads were washed three times with GPB, suspended in 500 µl GPB including 35S-labeled Coi12p recombinant proteins (1.2 to 2 µl of reaction of PURExpress *In Vitro* Protein Synthesis Kit), and incubated for ~2 h at 4°C.

The beads were washed five times with 0.1% BSA in GPB and boiled in 1x SDS-PAGE sample buffer. The samples were separated by SDS-PAGE. The gel was dried, wrapped in Saran wrap, and exposed overnight to a storage phosphor screen (GE Healthcare). Images were captured using a Typhoon Imager (GE Healthcare) and analyzed using ImageQuant software.

Identification of Coi12p-interacting proteins

Purification of Coi12p-associated proteins was performed using wild type B2086 and CU428 *or* COI12 *[*(B)-3-a-κ and (C)-2-b] strains and anti Coi12p antibodies cross-linked to AffiPrep beads (Bio-rad). Cross-linking was performed by incubating 150 µl of anti Coi12 serum and 100 µl beads in 20 mM DMP (dimethyl pilemidate), 0.2 M Sodium borate pH 9. Reaction was stopped by washing with 0.2 M ethanolamine, 250 mM Tris pH 8. To remove uncoupled heavy chains, beads were finally washed with 100 mM glycine pH 2 and directly used for immunoprecipitation.

Wild type B2086 and CU428 (5×10^7 cells) were starved, mated and collected at 3 hr post-mixing. Cells were lysed with one hundred strokes of a Dounce homogenizer in lysis buffer containing 50 mM Tris pH 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 1% Tween20, 0.2 mM PMSF, 1x complete proteinase inhibitor cocktail (Roche) and 1mM DTT. Soluble fractions were separated by centrifugation. Coi12p associated proteins were purified by immuno-affinity purification on AffiPrep beads (Bio-rad) crosslinked with anti-Coi12 antibody. Proteins were washed 3 times using lysis buffer and 3 times using lysis buffer without Tween20. Finally, Coi12p complexes were eluted with 0.2 M glycine-HCl pH 2.0 and neutralized using 2 M Tris-base. Purified material was separated by 8% SDS-PAGE and visualized by silver staining. To identify the Coi12p-associated protein, the eluate was directly analyzed by mass spectrometry.

Immunoprecipitation of Twi1p-scnRNA complex

To analyze Twi1p-associated scnRNAs, WT (CU428 and B2086) or *COI12KO [*(B)-3-a-κ and (C)-2-b] strains were mated. At 3 hr post-mixing, 5 x 10^6 cells from each cross were collected and were lysed with one hundred strokes of a Dounce homogenizer in lysis buffer containing 50 mM Tris pH 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 1% Tween 20, 0.2 mM PMSF, 1x complete proteinase inhibitor cocktail (Roche) and 0.4 U/ml RNasin (Promega). Soluble fractions were separated by centrifugation. 50 µl of this fraction was treated with TRIzol reagent (Invitrogen) and used as a source of input RNA. The remaining soluble fraction was incubated with anti-Twi1p agarose affinity gels (Sigma) over-night to precipitate Twi1p complexes. Twi1p-complexes were eluted by boiling the gels in 1x SDS-PAGE buffer. RNA in the elution was extracted using the TRIzol reagent (Invitrogen). scnRNA was separated on 15% denaturing polyacrylamide gels and stained with GelRed (Gentauer, Brussels, Belgium). Twi1p was detected by western blot using anti-Twi1p antibody (Aronica et al., 2008).

Analysis of scnRNAs

Total RNA was extracted from 8 x 10^5 starved or mating cells using the TRIzol reagent (Invitrogen), separated by electrophoresis in 15% denaturing polyacrylamide gels and stained with GelRed (Gentauer, Brussels, Belgium).

RdRP assay

Coi12p-associated complexes were immunoprecipitated as described above with a slight modification: anti-Coi12 antibody was not cross-linked to the beads and, before elution, Coi12p-complexes on beads were washed 3 times using lysis buffer [50 mM Tris pH 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 1% Tween20, 0.2 mM PMSF, 1x complete proteinase inhibitor cocktail (Roche) and 1mM DTT] and 3 times using washing buffer containing 50mM Hepes-NaOH pH 7.6, 20 mM ammonium acetate, 5mM MgCl2, 0.1 mM EDTA. Coi12p-complexes on beads were re-suspended in 100 µl washing buffer and 100 µl 80% glycerol, snap frozen and stored at -85ºC until the moment of use. For each reaction, 10 µl of beads were used. Template GFP RNA sense strand transcripts was synthesized by in vitro transcription (Ambion MAXI script kit) using template T7_EGFP_Sense DNA () and T7-EGFP-S-FW x EGFP-S-RV as primers. The RNA product was gel purified and processed to phenol/chloroform extraction. For testing primer-dependent enzymatic activity, equi-molar amount of primer EGFPAS1(UCACCUUCACCUGAAACAGA-AAAUUUGUG, annealing at position 75-100 of T7_EGFP_Sense RNA) was mixed with the template GFP RNA sense strand, denatured at 95 ºC for 5 min, and let re-nature slowly for about 1 hr in a pot of pre-heated water (almost boiling). 10-20 µM of template with or without annealed primer was used for reaction.

RdRP activity was tested in 10 µl reaction containing 50 mM HEPES-NaOH [pH 7.6], 20 mM ammonium acetate (NH4OAc), 5.6% (w/v) PEG8000, 5 mM MgCl2, 0.1 mM EDTA, 1 mM each of ATP and GTP, 0.2 mM each of CTP and UTP, 10-20 µM of ssRNA template, 0.8 unit/µl RNasin, 20 µCi ^{32}P-UTP, 10 µl of immunocomplex on beads. As negative control, one reaction without template RNA was set up. Reactions were incubated 2 hr at 30 ºC. The RNA product was processed to phenol/chloroform extraction and re-suspended in 40 µl DEPC-H_20 containing nuclease S1 (Fermentas) buffer. Each reaction was split into 4x 10 µl samples and treated with following concentrations of nuclease S1 (Fermentas), at 37 ºC, for 30 minutes: 0 u/µl, 0.1u/µl, 0.5u/µl, 5u/µl). 5 µl of each sample was re-suspended in 5 µl formamide and 1 µl loading dye, denatured at 95 °C for 5 min and loaded on 7% PAGE gel. The gel was run at 100-150 volt for about 1 hr at room temperature, dried and exposed overnight to a storage phosphor screen (GE Healthcare). Images were captured using a Typhoon Imager (GE Healthcare) and analyzed using ImageQuant software.

Electrophoretic Mobility Shift Assay

Electro-mobility shift assay (EMSA) protocol was adapted form Vargason et al. 2003. Association of Coi12p with 28 nt ds scnRNA with 2 nt, 3' overhangs, 28 nt ss scnRNA, 100 nt ssRNA (GFP RNA sense strand) or 28 nt DNA was tested. Synthetic 28 nt single strand sense scnRNA, and 28 nt DNA of the same sequence were purchased from Dharmacon and Life Technologies (Vienna, Austria) respectively. The 100 nt ssRNA was synthesized by in vitro transcription (Ambion MAXI script kit) in presence of 50 µCi [α-^{32}P] UTP. The template DNA for the in vitro transcription was prepared by PCR using pEGFP-neo4 and T7-EGFP-S-FW x EGFP-S-RV as primers. The RNA product was gel purified and processed to phenol/chloroform extraction. Single stranded 28 nt RNA and DNA oligonucleotides were radiolabeled in 50 pmol quanitities with 25 μCi [γ-^{32}P] ATP (Perkin Elmer Life Sciences) and 5 U of Optikinase (USB), according to the manufacturer's directions. For making 28 nt ds scnRNA, unlabled, 5'-phosphorylated antisense complementary strand was added to the inactivated and purified kinase reaction and the oligos were annealed at 90°C for 5 min and let slow cool down in a pot of boiling water. Coi12p was titrated in 30 µl binding reactions, against 0.3 nM radiolabeled RNA/DNA oligonucleotides in binding buffer containing 1M Tris-HCl (pH 7.5), 0.5M EDTA pH 8, 1 M DTT, 80% glycerol, 10% Tween20. Binding reactions were incubated at room temperature for 1 hr, adjusted to 7.5% (w/v) Ficoll 400, and 20 µl were analyzed by electrophoresis at a constant 100 V for 45 min through a 6% TBE DNA retardation gel (Novex, Invitrogen Inc.) in 0.5X TBE. The gels were then dried, exposed to a storage phosphor screen for 2–7 days, and bands corresponding to bound and free radiolabeled probe were quantified using a Typhoon Imager (GE Healthcare) and an ImageQuant software. The data from the direct binding experiments were fit to a rectangular hyperbola.

RT–PCR

Total RNA was extracted using TRIzol reagent (Invitrogen) from starved or conjugative wild-type (B2086 and CU428) cells. For cDNA synthesis, 5 µg of Turbo DNase (Ambion)-treated total RNA was retro-transcribed using 5 µl of Superscript II (Invitrogen) retrotrascriptase and 5 µl of random 6 mer primers. About 25 ng of cDNA were analysed by RT-PCR in 25-50 µl reaction, using primer specific for each *COI* gene or for *RPL21* (encoding L21 ribosomal protein) as a control.

Acknowledgments

I would like to thank Dr. Kazufumi Mochizuki for the opportunity to perform my PhD thesis in his lab, for all his support and enthusiasm as well as the constructive discussions with him. I am also very grateful to all members of Mochizuki Lab for the nice working atmosphere. I especially thank Dr. Kensuke Kataoka for providing *EGFP-COI12* strains and Tomoko Noto for technical advice. I also want to thank my PhD committee, Dr. Leonie Ringrose and Prof. Christian Seiser for guidance in my project. A final thanks goes to Maria Novatchkowa for the bioinformatics analysis of EST data, to Thomas Jenuwein for providing anti-H3 antibodies and to all the IMP/IMBA service facilities for making my everyday life easier. This project received funding from the European Research Council under the European Community's Sixth Framework Programme, from the Marie Curie Action ''Early Stage Training'' (grant agreement MEST-CT-2005-019676 EU FP6).

References

Allis CD, Chicoine LG, Richman R, Schulman IG. 1985. Deposition-related histone acetylation in micronuclei of conjugating Tetrahymena. *Proc Natl Acad Sci U S A 82*:8048-8052.

Allis CD, Richman R, Gorovsky MA, Ziegler YS, Touchstone B, Bradley WA, Cook RG. 1986. hv1 is an evolutionarily conserved H2A variant that is preferentially associated with active genes. *J Biol Chem 261*:1941-1948.

Allis-Jenuwein-Reinberg. 2006. Epigenetics. *CSHL Press*.

Apostol BL, Westaway SK, Abelson J, Greer CL. 1991. Deletion analysis of a multifunctional yeast tRNA ligase polypeptide. Identification of essential and dispensable functional domains. *J Biol Chem 266*:7445-7455.

Aravin AA, Hannon GJ, Brennecke J. 2007. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science 318*:761-764.

Aravin AA, Lagos-Quintana M, Yalcin A, Zavolan M, Marks D, Snyder B, Gaasterland T, Meyer J, Tuschl T. 2003. The small RNA profile during Drosophila melanogaster development. *Dev Cell 5*:337-350.

Aronica L, Bednenko J, Noto T, DeSouza LV, Siu KW, Loidl J, Pearlman RE, Gorovsky MA, Mochizuki K. 2008. Study of an RNA helicase implicates small RNA-noncoding RNA interactions in programmed DNA elimination in Tetrahymena. *Genes Dev 22*:2228-2241.

Axtell MJ, Snyder JA, Bartel DP. 2007. Common functions for diverse small RNAs of land plants. *Plant Cell 19*:1750-1769.

Bao N, Lye KW, Barton MK. 2004. MicroRNA binding sites in Arabidopsis class III HD-ZIP mRNAs are required for methylation of the template chromosome. *Dev Cell 7*:653-662.

Baulcombe D. 2005. RNA silencing. *Trends Biochem Sci 30*:290-293.

Baulcombe DC. 2007. Molecular biology. Amplified silencing. *Science 315*:199-200.

Bednenko J, Noto T, DeSouza LV, Siu KW, Pearlman RE, Mochizuki K, Gorovsky MA. 2009. Two GW repeat proteins interact with Tetrahymena thermophila argonaute and promote genome rearrangement. *Mol Cell Biol 29*:5020-5030.

Beermann S. 1977. The diminution of Heterochromatic chromosomal segments in Cyclops (Crustacea, Copepoda). *Chromosoma 60*:297-344.

Behrens SE, Tomei L, De Francesco R. 1996. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *Embo J 15*:12-22.

Bergstrom CT, McKittrick E, Antia R. 2003. Mathematical models of RNA silencing: unidirectional amplification limits accidental self-directed reactions. *Proc Natl Acad Sci U S A 100*:11511-11516.

Bernstein E, Caudy AA, Hammond SM, Hannon GJ. 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature 409*:363-366.

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell 128*:1089-1103.

Bressanelli S, Tomei L, Roussel A, Incitti I, Vitale RL, Mathieu M, De Francesco R, Rey FA. 1999. Crystal structure of the RNA-dependent

RNA polymerase of hepatitis C virus. *Proc Natl Acad Sci U S A 96*:13034-13039.

Brower-Toland B, Findley SD, Jiang L, Liu L, Yin H, Dus M, Zhou P, Elgin SC, Lin H. 2007. Drosophila PIWI associates with chromatin and interacts directly with HP1a. *Genes Dev 21*:2300-2311.

Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD. 1996. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell 84*:843-851.

Buhler M, Spies N, Bartel DP, Moazed D. 2008. TRAMP-mediated RNA surveillance prevents spurious entry of RNAs into the Schizosaccharomyces pombe siRNA pathway. *Nat Struct Mol Biol 15*:1015-1023.

Buker SM, Iida T, Buhler M, Villen J, Gygi SP, Nakayama J, Moazed D. 2007. Two different Argonaute complexes are required for siRNA generation and heterochromatin assembly in fission yeast. *Nat Struct Mol Biol 14*:200-207.

Butcher SJ, Grimes JM, Makeyev EV, Bamford DH, Stuart DI. 2001. A mechanism for initiating RNA-dependent RNA polymerization. *Nature 410*:235-240.

Callahan RC, Shalke G, Gorovsky MA. 1984. Developmental rearrangements associated with a single type of expressed alpha-tubulin gene in Tetrahymena. *Cell 36*:441-445.

Carmell MA, Xuan Z, Zhang MQ, Hannon GJ. 2002. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev 16*:2733-2742.

Cassidy-Hanley D, Bowen J, Lee JH, Cole E, VerPlank LA, Gaertig J, Gorovsky MA, Bruns PJ. 1997. Germline and somatic transformation of mating Tetrahymena thermophila by particle bombardment. *Genetics 146*:135-147.

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, Yao MC. 1996. Non-Mendelian, heritable blocks to DNA rearrangement are induced by loading the somatic nucleus of Tetrahymena thermophila with germ line-limited DNA. *Mol Cell Biol 16*:3658-3667.

Chalker DL, Yao MC. 2001. Nongenic, bidirectional transcription precedes and may promote developmental DNA deletion in Tetrahymena thermophila. *Genes Dev 15*:1287-1298.

Chatterjee S, Grosshans H. 2009. Active turnover modulates mature microRNA activity in Caenorhabditis elegans. *Nature 461*:546-549.

Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, Shiekhattar R. 2005. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature 436*:740-744.

Cherry JM, Blackburn EH. 1985. The internally located telomeric sequences in the germ-line chromosomes of Tetrahymena are at the ends of transposon-like elements. *Cell 43*:747-758.

Chung WJ, Okamura K, Martin R, Lai EC. 2008. Endogenous RNA interference provides a somatic defense against Drosophila transposons. *Curr Biol 18*:795-802.

Collins K, Gorovsky MA. 2005. Tetrahymena thermophila. *Curr Biol 15*:R317-318.

Cornell CT, Semler BL. 2002. Subdomain specific functions of the RNA polymerase region of poliovirus 3CD polypeptide. *Virology 298*:200-213.

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.

Coyne RS, Chalker DL, Yao MC. 1996. Genome downsizing during ciliate development: nuclear division of labor through chromosome restructuring. *Annu Rev Genet 30*:557-578.

Coyne RS, Nikiforov MA, Smothers JF, Allis CD, Yao MC. 1999. Parental expression of the chromodomain protein Pdd1p is required for completion of programmed DNA elimination and nuclear differentiation. *Mol Cell 4*:865-872.

Debeauchamp JL, Moses A, Noffsinger VJ, Ulrich DL, Job G, Kosinski AM, Partridge JF. 2008. Chp1-Tas3 interaction is required to recruit RITS to fission yeast centromeres and for maintenance of centromeric heterochromatin. *Mol Cell Biol 28*:2154-2166.

Deshpande G, Calhoun G, Schedl P. 2005. Drosophila argonaute-2 is required early in embryogenesis for the assembly of centric/centromeric heterochromatin, nuclear division, nuclear migration, and germ-cell formation. *Genes Dev 19*:1680-1685.

Duharcourt S, Yao MC. 2002. Role of histone deacetylation in developmentally programmed DNA rearrangements in Tetrahymena thermophila. *Eukaryot Cell 1*:293-303.

El-Shami M, Pontier D, Lahmy S, Braun L, Picart C, Vega D, Hakimi MA, Jacobsen SE, Cooke R, Lagrange T. 2007. Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. *Genes Dev 21*:2539-2544.

Filipowicz W, Jaskiewicz L, Kolb FA, Pillai RS. 2005. Post-transcriptional gene silencing by siRNAs and miRNAs. *Curr Opin Struct Biol 15*:331-341.

Fillingham JS, Thing TA, Vythilingum N, Keuroghlian A, Bruno D, Golding GB, Pearlman RE. 2004. A non-long terminal repeat retrotransposon family is restricted to the germ line micronucleus of the ciliated protozoan Tetrahymena thermophila. *Eukaryot Cell 3*:157-169.

Ghildiyal M, Seitz H, Horwich MD, Li C, Du T, Lee S, Xu J, Kittler EL, Zapp ML, Weng Z, Zamore PD. 2008. Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. *Science 320*:1077-1081.

Gillis AJ, Schuller AP, Skordalakes E. 2008. Structure of the Tribolium castaneum telomerase catalytic subunit TERT. *Nature 455*:633-637.

Goday C, Esteban MR. 2001. Chromosome elimination in sciarid flies. *Bioessays 23*:242-250.

Godiska R, James C, Yao MC. 1993. A distant 10-bp sequence specifies the boundaries of a programmed DNA deletion in Tetrahymena. *Genes Dev 7*:2357-2365.

Godiska R, Yao MC. 1990. A programmed site-specific DNA rearrangement in Tetrahymena thermophila requires flanking polypurine tracts. *Cell 61*:1237-1246.

Gonzalez S, Pisano DG, Serrano M. 2008. Mechanistic principles of chromatin remodeling guided by siRNAs and miRNAs. *Cell Cycle 7*:2601-2608.

Gregory RI, Chendrimada TP, Cooch N, Shiekhattar R. 2005. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell 123*:631-640.

Grishok A. 2005. RNAi mechanisms in Caenorhabditis elegans. *FEBS Lett 579*:5932-5939.

Hall IM, Shankaranarayana GD, Noma K, Ayoub N, Cohen A, Grewal SI. 2002. Establishment and maintenance of a heterochromatin domain. *Science 297*:2232-2237.

Hamilton AJ, Baulcombe DC. 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science 286*:950-952.

Hammond SM, Bernstein E, Beach D, Hannon GJ. 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature 404*:293-296.

Hayashi T, Hayashi H, Fusauchi Y, Iwai K. 1984. Tetrahymena histone H3. Purification and two variant sequences. *J Biochem 95*:1741-1749.

Haynes KA, Caudy AA, Collins L, Elgin SC. 2006. Element 1360 and RNAi components contribute to HP1-dependent silencing of a pericentric reporter. *Curr Biol 16*:2222-2227.

Henriette M. Kurth and KK u.

Herr AJ, Jensen MB, Dalmay T, Baulcombe DC. 2005. RNA polymerase IV directs silencing of endogenous DNA. *Science 308*:118-120.

Jacob F, Monod J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol 3*:318-356.

Karrer KM. 2000. Tetrahymena genetics: two nuclei are better than one. *Methods Cell Biol 62*:127-186.

Kawamura Y, Saito K, Kin T, Ono Y, Asai K, Sunohara T, Okada TN, Siomi MC, Siomi H. 2008. Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature 453*:793-797.

Kim DH, Saetrom P, Snove O, Jr., Rossi JJ. 2008. MicroRNA-directed transcriptional gene silencing in mammalian cells. *Proc Natl Acad Sci U S A 105*:16230-16235.

King BO, Yao MC. 1982. Tandemly repeated hexanucleotide at Tetrahymena rDNA free end is generated from a single copy during development. *Cell 31*:177-182.

Kubota S, Kuro-o M, Mizuno S, Kohno S. 1993. Germ line-restricted, highly repeated DNA sequences and their chromosomal localization in a Japanese hagfish (Eptatretus okinoseanus). *Chromosoma 102*:163-173.

LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D. 2005. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell 121*:713-724.

Lau NC, Lim LP, Weinstein EG, Bartel DP. 2001. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. *Science 294*:858-862.

Laurila MR, Makeyev EV, Bamford DH. 2002. Bacteriophage phi 6 RNA-dependent RNA polymerase: molecular details of initiating nucleic acid synthesis without primer. *J Biol Chem 277*:17117-17124.

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.

Li CF, Pontes O, El-Shami M, Henderson IR, Bernatavichute YV, Chan SW, Lagrange T, Pikaard CS, Jacobsen SE. 2006. An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in Arabidopsis thaliana. *Cell 126*:93-106.

Li HW, Ding SW. 2005. Antiviral silencing in animals. *FEBS Lett 579*:5965-5973.

Lipardi C, Baek HJ, Wei Q, Paterson BM. 2005. Analysis of short interfering RNA function in RNA interference by using Drosophila embryo extracts and schneider cells. *Methods Enzymol 392*:351-371.

Liu Y, Mochizuki K, Gorovsky MA. 2004. Histone H3 lysine 9 methylation is required for DNA elimination in developing macronuclei in Tetrahymena. *Proc Natl Acad Sci U S A 101*:1679-1684.

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.

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.

Maida Y, Yasukawa M, Furuuchi M, Lassmann T, Possemato R, Okamoto N, Kasim V, Hayashizaki Y, Hahn WC, Masutomi K. 2009. An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA. *Nature 461*:230-235.

Maine EM, Hauth J, Ratliff T, Vought VE, She X, Kelly WG. 2005. EGO-1, a putative RNA-dependent RNA polymerase, is required for heterochromatin assembly on unpaired dna during C. elegans meiosis. *Curr Biol 15*:1972-1978.

Makeyev EV, Bamford DH. 2002. Cellular RNA-dependent RNA polymerase involved in posttranscriptional gene silencing has two distinct activity modes. *Mol Cell 10*:1417-1427.

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.

Martindale DW, Allis CD, Bruns PJ. 1985. RNA and protein synthesis during meiotic prophase in Tetrahymena thermophila. *J Protozool 32*:644-649.

Martins A, Shuman S. 2004. Characterization of a baculovirus enzyme with RNA ligase, polynucleotide 5'-kinase, and polynucleotide 3'-phosphatase activities. *J Biol Chem 279*:18220-18231.

Miao W, Xiong J, Bowen J, Wang W, Liu Y, Braguinets O, Grigull J, Pearlman RE, Orias E, Gorovsky MA. 2009. Microarray analyses of gene expression during the Tetrahymena thermophila life cycle. *PLoS One 4*:e4429.

Moazed D. 2009. Small RNAs in transcriptional gene silencing and genome defence. *Nature 457*:413-420.

Mochizuki K. 2008. High efficiency transformation of Tetrahymena using a codon-optimized neomycin resistance gene. *Gene 425*:79-83.

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.

Mochizuki K, Gorovsky MA. 2004a. Conjugation-specific small RNAs in Tetrahymena have predicted properties of scan (scn) RNAs involved in genome rearrangement. *Genes Dev 18*:2068-2073.

Mochizuki K, Gorovsky MA. 2004b. Small RNAs in genome rearrangement in Tetrahymena. *Curr Opin Genet Dev 14*:181-187.

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, Novatchkova M, Loidl J. 2008. DNA double-strand breaks, but not crossovers, are required for the reorganization of meiotic nuclei in Tetrahymena. *J Cell Sci 121*:2148-2158.

Morris KV, Chan SW, Jacobsen SE, Looney DJ. 2004. Small interfering RNA-induced transcriptional gene silencing in human cells. *Science 305*:1289-1292.

Motamedi MR, Verdel A, Colmenares SU, Gerber SA, Gygi SP, Moazed D. 2004. Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell 119*:789-802.

Nikiforov MA, Gorovsky MA, Allis CD. 2000. A novel chromodomain protein, pdd3p, associates with internal eliminated sequences during macronuclear development in Tetrahymena thermophila. *Mol Cell Biol 20*:4128-4134.

Noma K, Sugiyama T, Cam H, Verdel A, Zofall M, Jia S, Moazed D, Grewal SI. 2004. RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. *Nat Genet 36*:1174-1180.

Noto a. In press. The Tetrahymena Argonaute-binding protein Giw1p directs a mature Argonaute-siRNA complex to the nucleus. *Cell*.

Okamura K, Ishizuka A, Siomi H, Siomi MC. 2004. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev 18*:1655-1666.

Pak J, Fire A. 2007. Distinct populations of primary and secondary effectors during RNAi in C. elegans. *Science 315*:241-244.

Pal-Bhadra M, Leibovitch BA, Gandhi SG, Rao M, Bhadra U, Birchler JA, Elgin SC. 2004. Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. *Science 303*:669-672.

Patil NS, Hempen PM, Udani RA, Karrer KM. 1997. Alternate junctions and microheterogeneity of Tlr1, a developmentally regulated DNA rearrangement in Tetrahymena thermophila. *J Eukaryot Microbiol 44*:518-522.

Plasterk RH. 2002. RNA silencing: the genome's immune system. *Science 296*:1263-1265.

Pontes O, Li CF, Nunes PC, Haag J, Ream T, Vitins A, Jacobsen SE, Pikaard CS. 2006. The Arabidopsis chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. *Cell 126*:79-92.

Prescott DM. 1994. The DNA of ciliated protozoa. *Microbiol Rev 58*:233-267.

Reinhart BJ, Bartel DP. 2002. Small RNAs correspond to centromere heterochromatic repeats. *Science 297*:1831.

Reiss D, Josse T, Anxolabehere D, Ronsseray S. 2004. aubergine mutations in Drosophila melanogaster impair P cytotype determination by telomeric P elements inserted in heterochromatin. *Mol Genet Genomics 272*:336-343.

Rexer CH, Chalker DL. 2007. Lia1p, a novel protein required during nuclear differentiation for genome-wide DNA rearrangements in Tetrahymena thermophila. *Eukaryot Cell 6*:1320-1329.

Ruby JG, Jan CH, Bartel DP. 2007. Intronic microRNA precursors that bypass Drosha processing. *Nature 448*:83-86.

Sarot E, Payen-Groschene G, Bucheton A, Pelisson A. 2004. Evidence for a piwi-dependent RNA silencing of the gypsy endogenous retrovirus by the Drosophila melanogaster flamenco gene. *Genetics 166*:1313-1321.

Saveliev SV, Cox MM. 1996. Developmentally programmed DNA deletion in Tetrahymena thermophila by a transposition-like reaction pathway. *Embo J 15*:2858-2869.

Seto AG, Kingston RE, Lau NC. 2007. The coming of age for Piwi proteins. *Mol Cell 26*:603-609.

Shiu PK, Raju NB, Zickler D, Metzenberg RL. 2001. Meiotic silencing by unpaired DNA. *Cell 107*:905-916.

Sijen T, Fleenor J, Simmer F, Thijssen KL, Parrish S, Timmons L, Plasterk RH, Fire A. 2001. On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell 107*:465-476.

Sijen T, Steiner FA, Thijssen KL, Plasterk RH. 2007. Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. *Science 315*:244-247.

Smothers JF, Madireddi MT, Warner FD, Allis CD. 1997. Programmed DNA degradation and nucleolar biogenesis occur in distinct organelles during macronuclear development in Tetrahymena. *J Eukaryot Microbiol 44*:79-88.

Song JJ, Joshua-Tor L. 2006. Argonaute and RNA--getting into the groove. *Curr Opin Struct Biol 16*:5-11.

Sugiyama T, Cam H, Verdel A, Moazed D, Grewal SI. 2005. RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proc Natl Acad Sci U S A 102*:152-157.

Taverna SD, Coyne RS, Allis CD. 2002. Methylation of histone h3 at lysine 9 targets programmed DNA elimination in tetrahymena. *Cell 110*:701-711.

Till S, Lejeune E, Thermann R, Bortfeld M, Hothorn M, Enderle D, Heinrich C, Hentze MW, Ladurner AG. 2007. A conserved motif in Argonaute-interacting proteins mediates functional interactions through the Argonaute PIWI domain. *Nat Struct Mol Biol 14*:897-903.

Ting AH, Schuebel KE, Herman JG, Baylin SB. 2005. Short double-stranded RNA induces transcriptional gene silencing in human cancer cells in the absence of DNA methylation. *Nat Genet 37*:906-910.

Tomari Y, Zamore PD. 2005. Perspective: machines for RNAi. *Genes Dev 19*:517-529.

Vazquez F, Vaucheret H, Rajagopalan R, Lepers C, Gasciolli V, Mallory AC, Hilbert JL, Bartel DP, Crete P. 2004. Endogenous trans-acting siRNAs regulate the accumulation of Arabidopsis mRNAs. *Mol Cell 16*:69-79.

Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SI, Moazed D. 2004. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science 303*:672-676.

Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science 297*:1833-1837.

Wang HW, Noland C, Siridechadilok B, Taylor DW, Ma E, Felderer K, Doudna JA, Nogales E. 2009. Structural insights into RNA processing by the human RISC-loading complex. *Nat Struct Mol Biol 16*:1148-1153.

Wassenegger M, Krczal G. 2006. Nomenclature and functions of RNA-directed RNA polymerases. *Trends Plant Sci 11*:142-151.

Wuitschick JD, Gershan JA, Lochowicz AJ, Li S, Karrer KM. 2002. A novel family of mobile genetic elements is limited to the germline genome in Tetrahymena thermophila. *Nucleic Acids Res 30*:2524-2537.

Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, Boulay J, Regnault B, Devaux F, Namane A, Seraphin B, Libri D, Jacquier A.

2005. Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell 121*:725-737.

Xu Q, Teplow D, Lee TD, Abelson J. 1990. Domain structure in yeast tRNA ligase. *Biochemistry 29*:6132-6138.

Yao MC. 1981. Ribosomal RNA gene amplification in Tetrahymena may be associated with chromosome breakage and DNA elimination. *Cell 24*:765-774.

Yao MC, Chao JL. 2005. RNA-guided DNA deletion in Tetrahymena: an RNAi-based mechanism for programmed genome rearrangements. *Annu Rev Genet 39*:537-559.

Yao MC, Fuller P, Xi X. 2003. Programmed DNA deletion as an RNA-guided system of genome defense. *Science 300*:1581-1584.

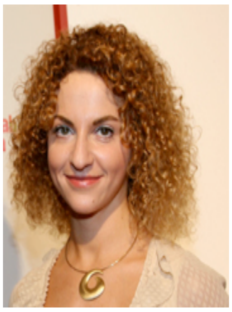
Yao MC, Gorovsky MA. 1974. Comparison of the sequences of macro- and micronuclear DNA of Tetrahymena pyriformis. *Chromosoma 48*:1-18.

Yao MC, Yao CH, Halasz LM, Fuller P, Rexer CH, Wang SH, Jain R, Coyne RS, Chalker DL. 2007. Identification of novel chromatin-associated proteins involved in programmed genome rearrangements in Tetrahymena. *J Cell Sci 120*:1978-1989.

Yao MC, Yao CH, Monks B. 1990. The controlling sequence for site-specific chromosome breakage in Tetrahymena. *Cell 63*:763-772.

Zamore PD, Tuschl T, Sharp PA, Bartel DP. 2000. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell 101*:25-33.

Curriculum Vitae



Personal information

First name(s) / Surname(s)	**Lucia Aronica**		
Telephone(s)	(+39-81) 7518925	Mobile	(+43-650) 4562030
E-mail(s)	luciaaronica@tiscali.it		
Nationality	Italian		
Date of birth	19 Oct 81		

Education and training

Dates	Since Jul 2006
Title of qualification awarded	Ph. D. study
Principal subjects / occupational skills covered	Molecular Biology- RNA interference
Name and type of organisation providing education and training	IMBA (Institute of Molecular Biotechnology of the Austrian Academy of Science), Vienna (Austria) (University of Vienna)
Level in national or international classification	ISCED 6
Dates	2000 - 2005
Title of qualification awarded	Laurea in Medical Biotechnology, 110/110 cum laude with special mention
Principal subjects / occupational skills covered	Molecular Biology, Genetics, Chemistry and Biochemistry, Pharmacology, Anatomy and Physiology
Name and type of organisation providing education and training	Università degli Studi Federico II, Naples (Italy)
Level in national or international classification	ISCED 5
Dates	1995 - 2000
Title of qualification awarded	Diploma di Maturità Classica, Grade: 100/100
Principal subjects / occupational skills covered	History, Philosophy, Latin, ancient Greek, Italian Literature
Name and type of organisation providing education and training	Liceo Classico Statale Giuseppe Garibaldi, Naples (Italy)
Level in national or international classification	ISCED 3

Dates	04/2008
Title of qualification awarded	Certificate of attendance
Principal subjects / occupational skills covered	Project Management and Intellectual Property Rights
Name and type of organisation providing education and training	University of Vienna, Vienna (Austria)
Dates	04/2008
Title of qualification awarded	Certificate of attendance
Principal subjects / occupational skills covered	European Patent Office workshop (Patent law, Intellectual property rights)
Name and type of organisation providing education and training	European Patent Office, Vienna (Austria)
Dates	08/2008
Title of qualification awarded	Certificate of attendance
Principal subjects / occupational skills covered	Presentation skill course
Name and type of organisation providing education and training	IMBA GmbH, Vienna (Austria)
Dates	10/2007
Title of qualification awarded	Certificate of attendance
Principal subjects / occupational skills covered	Business in Biotechnology
Name and type of organisation providing education and training	Vienna Biocenter, Vienna (Austria)
Dates	08/2005
Title of qualification awarded	Certificate of attendance
Principal subjects / occupational skills covered	Biomedical Engineering
Name and type of organisation providing education and training	New University of Lisbon, Almada (Portugal)
Dates	04/2005
Title of qualification awarded	Certificate of attendance
Principal subjects / occupational skills covered	Alternative energy sources
Name and type of organisation providing education and training	Institute for Human Factors and Technology Management, Stuttgart (Germany)
Dates	02/2005
Title of qualification awarded	Certificate of attendance
Principal subjects / occupational skills covered	Telemedicine
Name and type of organisation providing education and training	University of Sophia, Sophia (Bulgaria)

Personal skills and competences

Mother tongue(s) **Italian**

Other language(s)

Self-assessment	Understanding				Speaking				Writing	
European level (*)	Listening		Reading		Spoken interaction		Spoken production			
English	C2	Proficient user	C2	Proficient user	C2	Proficient user	C2	Proficient user	C1	Proficient user
German	C1	Proficient user	C1	Proficient user	C1	Proficient user	C1	Proficient user	B2	Independent us

(*) *Common European Framework of Reference (CEF) level*

Social skills and competences

Excellent presentation skills, gained through public speaking at meetings and specific courses and awarded in specific competitons (Famelab Austria, Famelab international)

Technical skills and competences

Scientific expertise. Proficiency in English and very good knowledge of German language. Basic knowledge in business administration, marketing, project management, intellectual property rights and patent examination procedure (see courses listed above)

Computer skills and competences

Word processing, spreadsheet programs, presentations (MS Power-Point), graphics (Photoshop), database searching

Additional information

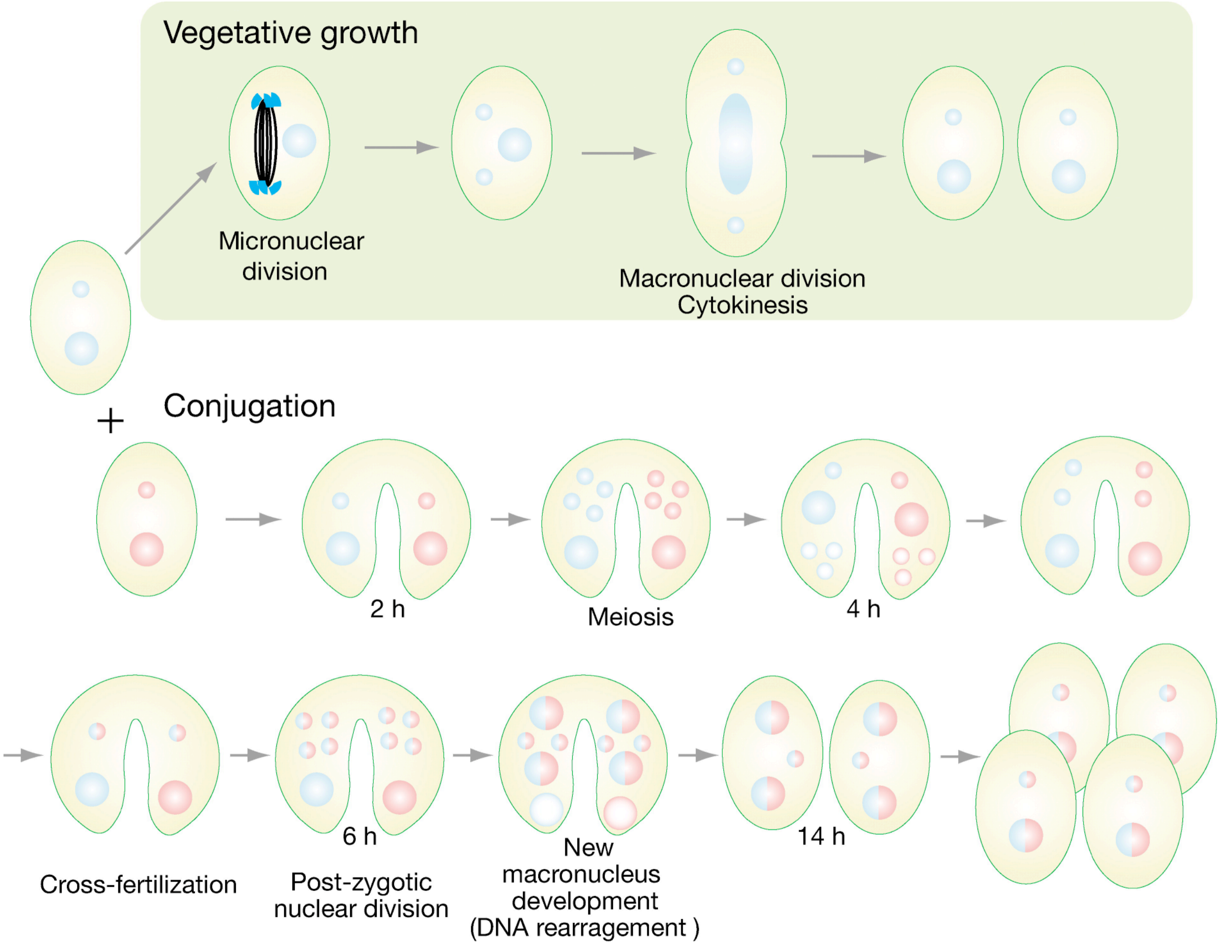
PUBLICATIONS:

Genes and Development, August 15, 2008 22:2228-2241. Aronica L, Bednenko J*, Noto T*, DeSouza LV, Siu KW, Loidl J, Pearlman RE, Gorovsky MA, Mochizuki K. “Study of an RNA helicase implicates small RNA–noncoding RNA interactions in programmed DNA elimination in Tetrahymena”

* Equally contributing authors

AWARDS:

Winner of the scientific communication competition "Famelab Austria 2009" supported by the British Council (jury price and audience price), Vienna, Austria. Third place in "Famelab international 2009". Italian finalist at the VIII International Olympics of Philosophy 2000, Franz-Hitze-Haus – Muenster, Germany.



Yao M-C, Chao J-L. 2005.
Annu. Rev. Genet. 39:537–59

Figure 1.Life cycle of Tetrahymena.
The top panel shows the vegetative growth of Tetrahymena, which occurs when nutrients are plentiful. The bottom panel shows the conjugation process between two Tetrahymenacells of different mating types in starvation conditions (Yao M-C and Chao J-L, 2005) .

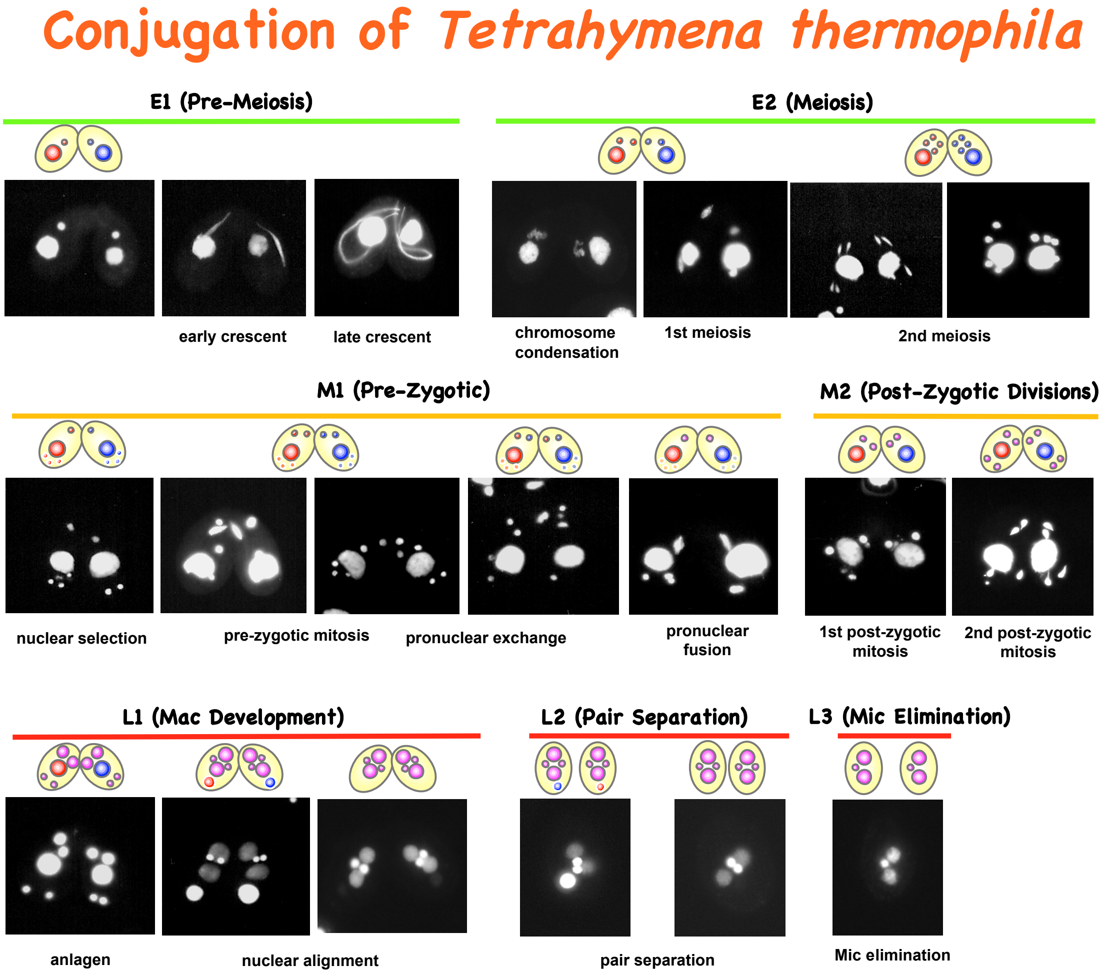
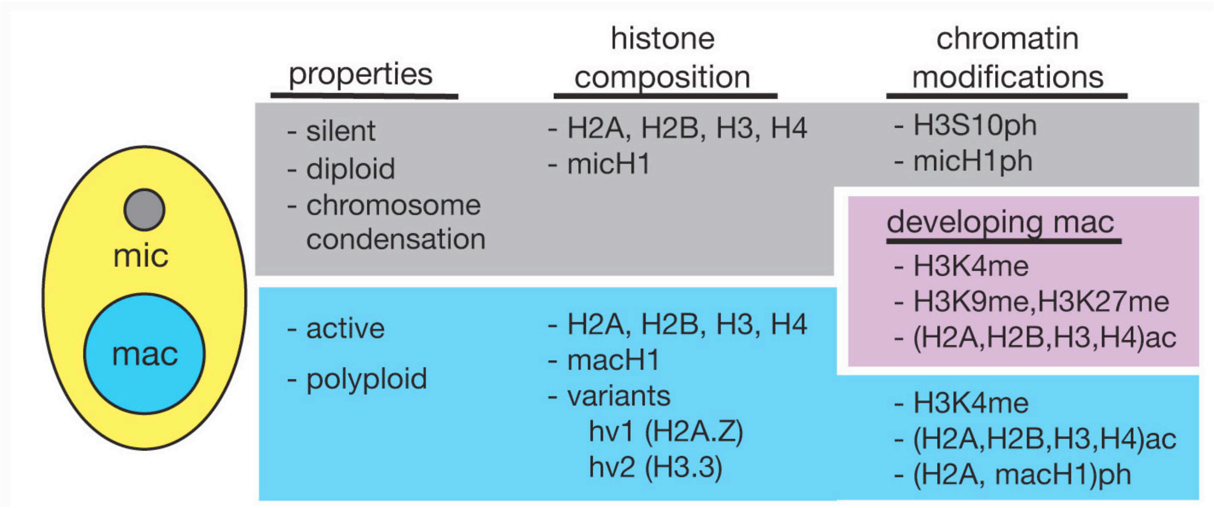


Figure 2. Nuclear events of conjugation in *Tetrahymena thermophila*. Schematic drawings and DAPI stained images of cells for each cytological stage of conjugation are shown. In this study, we distinguish 7 conjugative stages:E1 (occurs at approximately 0.5-3 hr post-mixing), E2 (3-4 hr), M1 (4-5 hr), M2 (5-6hr), L1 (6-10 hr), L2 (10-12 hr) and L3 (12 hr~), as shown in the figure. See Martindale et al. (1982) and Karrer (2000) for more detailed description of *Tetrahymena* conjugation.



	properties	histone composition	chromatin modifications
mic	- silent - diploid - chromosome condensation	- H2A, H2B, H3, H4 - micH1	- H3S10ph - micH1ph
developing mac			- H3K4me - H3K9me,H3K27me - (H2A,H2B,H3,H4)ac
mac	- active - polyploid	- H2A, H2B, H3, H4 - macH1 - variants hv1 (H2A.Z) hv2 (H3.3)	- H3K4me - (H2A,H2B,H3,H4)ac - (H2A, macH1)ph

Epigenetics © 2006 Cold Spring Harbor Laboratory Press

Figure 3. Nuclear dymorphism of *Tertahymena*.
The germline micronucleus, the somatic macronucleus and the new developing macronucleus contain different hystone complements and modifications. Those known to occur specifically in one of these genome are listed (Epigenetics ©, 2006 Cold Spring Harbor Laboratory Press).

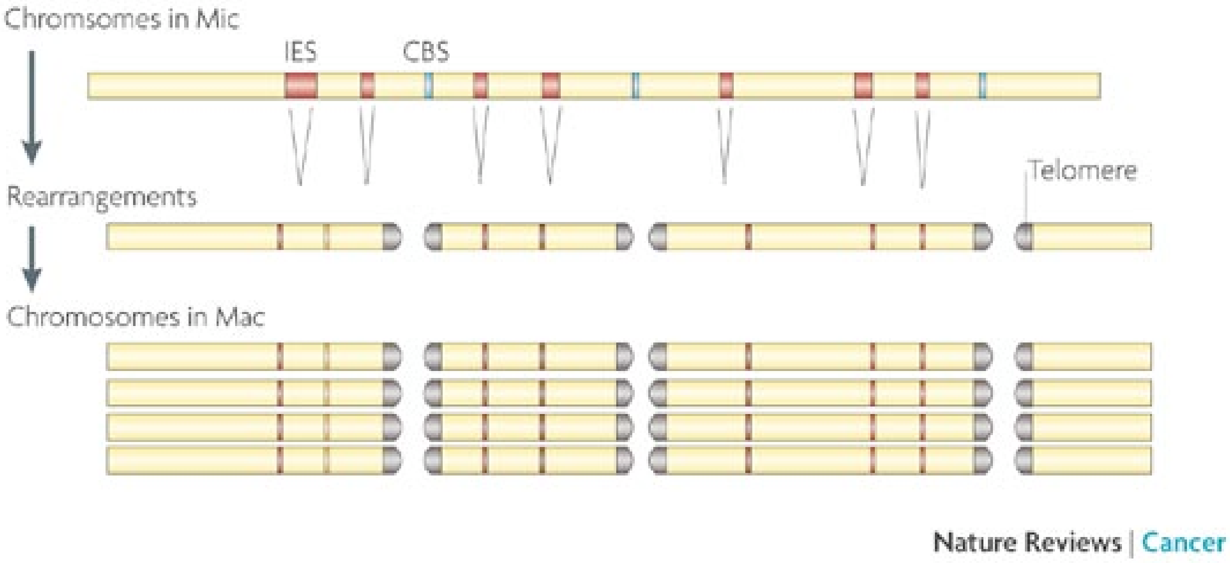


Figure 4. DNA rearrangements in *Tetrahymena thermophila*
DNA rearrangements (step 6) involve two major processes. One process is the deletion of internal eliminated sequences (IESs, also known as deletion elements). Each IES ranges from several hundred base pairs to >10 kb in size. There are more than 6,000 IESs in the micronuclear genome and the deletion processes collectively eliminates 15% of the genome. The second process is the formation of chromosome breaks at chromosome breakage sequences (CBS) and the de novo addition of telomeres (grey). This process fragments five germline chromosomes in micronuclei into 200–300 mini chromosomes that are then endo-replicated to complete the polyploidy (45n) of the new macronucleus (Tanaka and Yao, 2009).

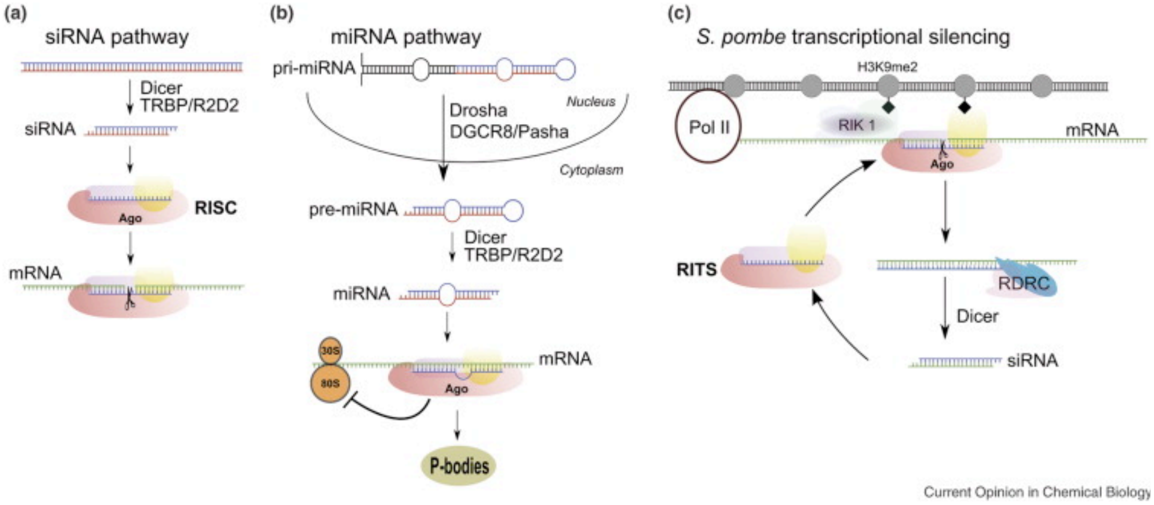


Figure 5. RNA interference pathways involving the Argonaute protein subfamily. A) In the siRNA pathway, exogenous dsRNAs are processed into siRNAs by Dicer complexes. The siRNAs are loaded into Argonaute proteins that are at the centre of the RNA-induced silencing complex (RISC) and catalyze endonucleolytic cleavage of siRNA-complementary mRNA , a process called slicing. B) MicroRNAs (miRNAs) are derived from endogenous hairpin precursors transcribed in the nucleus (pri-miRNA). Drosha containing complexes process the pri-miRNA into pre-miRNA that is transported to the cytoplasm where mature miRNAs are made by Dicer. Argonaute is loaded with miRNAs in miRNP complexes and is guided to the 3’ UTR of mature mRNA transcripts. miRNPs are believed to not slice, but instead inhibit translation or sequester transcripts to cytoplasmic foci called P-bodies. C) Transcriptional silencing and spreading in *S. pombe* is dependent on the RITS complex that contains Argonaute, Tas3 and Chp1 proteins. Silencing is achieved through the formation of heterochromatin induced by H3 Lys9 dimethylation (H3K9me2, black diamonds). An siRNA-guided RITS complex slices nascent transcripts synthesized by Pol II. RDRC converts the cleaved transcript into a dsRNA substrate for Dicer that produces secondary siRNAs. Loading into new RITS complexes for additional targeting is coupled to the Rik1 complex, which contains a histone methyltransferase resulting in H3K9me2 and spreading of silencing (Faehnle and Joshua-Tor, *Curr Opin Chem Biol*. 2007).

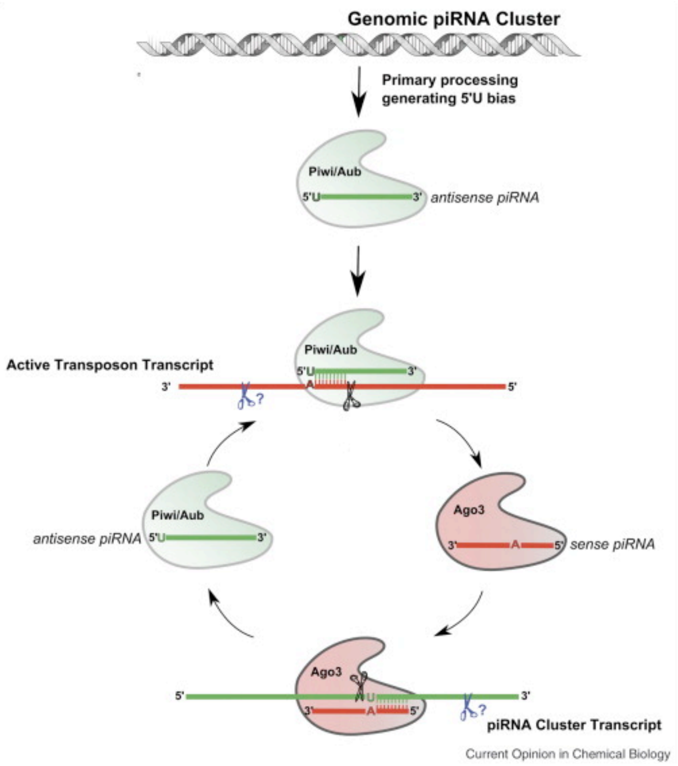


Figure 6. piRNA biogenesis mechanism.
Antisense Piwi-interacting RNAs (piRNAs) in flies associate with Piwi and Aubergine (Aub) proteins. A 10 bp complementary relationship exists for antisense Piwi/Aub piRNAs with sense-oriented Ago3 piRNAs; therefore, Slicer cleavage of active transposon transcripts by Piwi/Aub is proposed to generate the 5’ end of sense piRNAs bound to Ago3. This process is coupled with an unknown endonuclease to generate the 3’ end of mature sense piRNA. Sense piRNAs (10 bp complementary to antisense piRNAs) bound to Ago3 target piRNA cluster transcripts that would initiate the biogenesis of antisense Piwi/Aub piRNAs (Faehnle and Joshua-Tor*Curr Opin Chem Biol*. 2007).

.

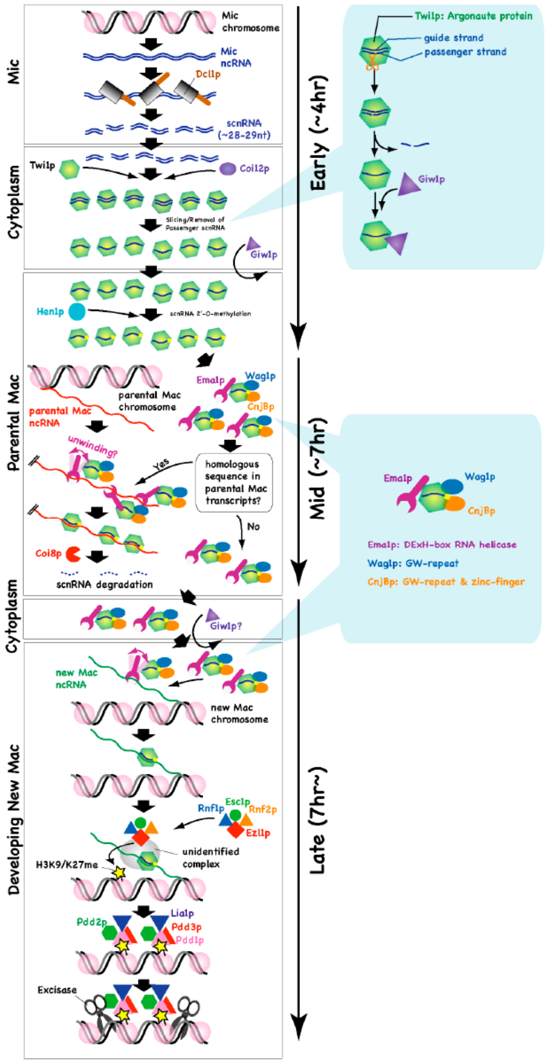
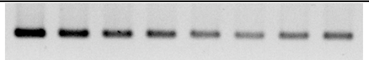
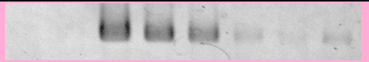
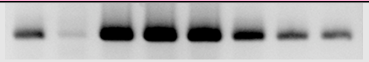
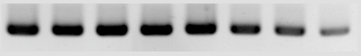
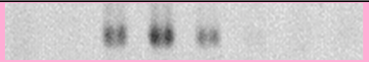
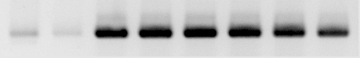
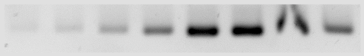
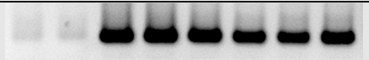
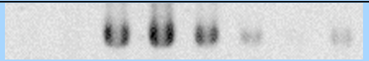
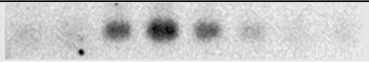
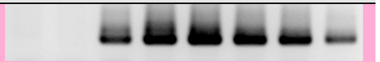
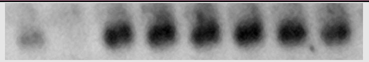
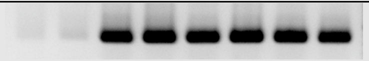
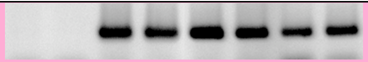
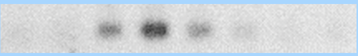
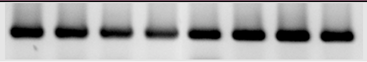
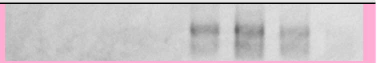
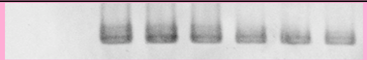
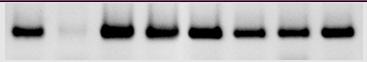
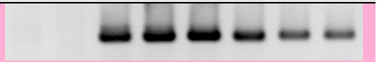
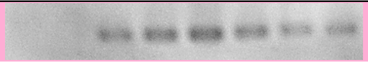
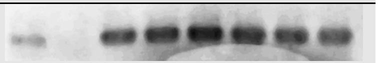
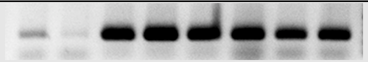
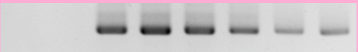
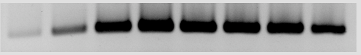
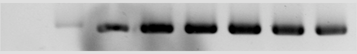
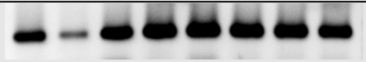
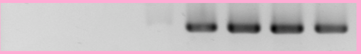
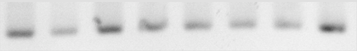
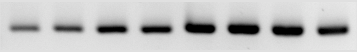
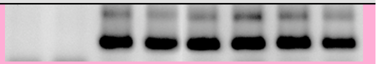
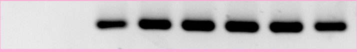
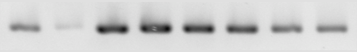
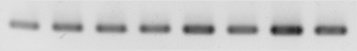
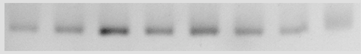
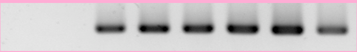
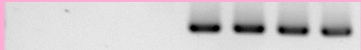
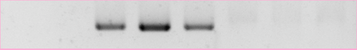
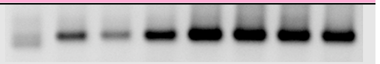


Figure 7. The scnRNA model.

In the early stages of conjugation (soon after mating, in the first 4 hr post-mixing), bidirectional ncRNA transcription occurs in the Mic. Double stranded Mic ncRNAs (shown as blue wavy lines in Fig. 7) are processed to ~28 nt scnRNAs by Dcl1p. Then, scnRNAs are transferred to the cytoplasm where they form complex with the Argonaute Twi1p. Twi1p cuts by endonuclease cleavage (slicing) the passenger scnRNA strand, which is removed to produce a mature Twi1p-scnRNA complex specifically recognized by Giw1p. Twi1p-Giw1p complex is transferred into the Mac. In parallel, ncRNAs are made from parental Mac chromosomes (red wavy lines) and homologous scnRNAs base-pair with them and are selectively degraded. Ema1p (drawn in purple) likely functions in this selective elimination of scnRNAs, possibly by unwinding the parental Mac ncRNAs to enhance the scnRNA-ncRNAs interaction. Then, Ema1p and two additional novel proteins, CnjBp (a Tex domain, GW repeat and CCHC zinc-finger containing protein) and Wag1p (a GW repeat protein) (Bednenko et al., 2009) transfer IES-specific scnRNA to the developing new Mac in the late stages of conjugation (>7 hr) and recruit a Policomb-like complex containing a histone methyltransferase, Ezl1p (Liu et al., 2007), two zink-finger protein,Rnf1 and Rnf2, and one ESC-homolog, Esc1p. This complex induces H3K9/K27me on the chromatin transcribing homologous ncRNAs (green wavy lines). Then, the chromodomain protein Pdd1p binds methylated IESs promoting the formation of heterochromatin-like structures, referred as dumposome (Madireddi et al., 1996). Finally, an unidentified endonuclease (excisase) eventually removes IESs and rejoins the flanking macronuclear destined sequences.

Gene name	TTHREM # GenBank #	Old annotation	# EST	Homology/ domain	RT-PCR vegetative conjugation
					E S 2 4 6 8 10 12
RPL21	Loading control for RT-PCR				
TWI1	01161040 XM_001018296		319	PAZ and Piwi domains	
PDD1	00125280 XM_001016207		148	chromodomain	
CNJB	01091290 XM_001025435		91	GW repear, zinc-finger	
COI1 (GIW1)	01276320 XM_001029843	304.m00015	62		
	00155590 XM_001009685	13.m00306	52	HMG-box family	
	00649180 XM_001032312	94.m00150	43		
COI2	00046930 XM_001014688	3.m02016	41	-	
	00572190 XM_001027725	76.m00251	40		
COI3	00526270 XM_001028038	66.m00164	39 (24)		
			39 (15)	Peptidase family C54	
	00219320 XM_001020604	18.m00386	38		
PDD2	00283530 XM_001018209	53.m00256	37		
DMC1	00459230 XM_001024231	12.m00332	37	Rad51/Dmc1	
	00148990 XM_001021534	65.m00215	35		
COI4	00522820 XM_001014460	57.m00176	32	-	
	00474360 XM_001023920	149.m00053	31		
LIA6	00849260 XM_001019228	5.m00549	30	Partially similar to *COI16*	

Gene name	TTHREM # GenBank #	Old annotation	# EST	Homology/ domain	RT-PCR vegetative: E S conjugation: 2 4 6 8 10 12
COI5	00079530 XM_001015793	53.m00256	29	Cyclin, N-terminal domain	
MND1A	00300660 XM_001024593	28.m00320	28	Mnd1-like	
PDD3	01109990 XM_001030208		28	chromodomain	
COI6	00086720 XM_001012720	6.m00495	28	chromodomain	
	00439170 XM_001017814	49.m00246	28		
COI7	00348580 XM_001023013	36.m00234	27	-	
COI8	00134800 XM_001019660	10.m00353	26	-	
	00268140 XM_001015949	24.m00280	26		
COI9	00531890 XM_001024314	68.m00138	25	Zn-finger, RING	
COI10	01367700 XM_001026268	350.m00014	25	-	
TWI11	00144830 XM_001011123	11.m00388	25	PAZ and Piwi domains	
COI11	00189440 XM_001016620	15.m00405	25	Leucine rich repeat	
	00732800 XM_001023678	119.m00126	25		
	00185640 XM_001008835	14.m00453	24	jmjC domain	
	00300650 XM_001024592	28.m00319	24		
RPB2	00077230 XM_001015664		24	Rpb2p	
COI14	00420400 XM_001033295	46.m00220	24	Zinc-finger domain?	
	01358410 XM_001013522	345.m00014	24	ABC transporter	

Gene name	TTHREM # GenBank #	Old annotation	# EST	Homology/ domain	RT-PCR
					vegetative E S conjugation 2 4 6 8 10 12
EMA1	00088150 XM_001012764	6.m00539	24	DExH-box RNA helicase	
	00439300 XM_001017827	49.m00259	24		
	00433790 XM_001011404	48.m00283	24	Tetratricopeptide repeat	
	00361840 XM_001007394	38.m00234	23	TFIIS?	
COI15	00006160 XM_001008097	1.m00672	23	-	
	00128920 XM_001016373	9.m00477	23		
COI16	00564480 XM_001022047	75.m00184	23	Partially similar to *LIA6* (00849260)	
	00412010 XM_001020895	44.m00286	23		
COI12	00402050 XM_001014078	62.m00216	22	-	
COI17	00497670 XM_001027955	43.m00267	22	-	
	00133710 XM_001019650	10.m00343	22		
	00034970 XM_971905	2.m02396	22	Protein kinase	
	00501010 XM_001022260	63.m00223	21		
	00238970 XM_001024822	20.m00300	21		
COI18	00197670 XM_972205	17.m00242	20	-	
COI19	01085480 XM_001030381	226.m00036	19		
COI20	00049220 XM_001014817	3.m01752	19		
	00664050 XM_001020171	98.m00195	19		

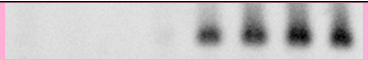
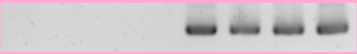
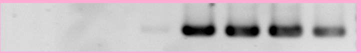
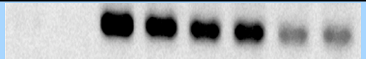
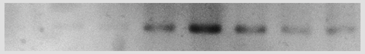
Gene name	TTHREM # GenBank #	Old annotation	# EST	Homology/ domain	RT-PCR vegetative conjugation E S 2 4 6 8 10 12
COI13	00585190 XM_001032624	80.m00171	19	chromodomain	
LIA1	00675900 XM_001015062	101.m00186	19		
COI21	00193970 XM_001017150	16.m00337	19	lipase	
COI22	00237610 XM_001024791	20.m00269	19		
EZL1	00335780 NW_001219307	34.m00260	19	SET domain	
	00112710 XM_001010666	8.m00345	19		

Figure 8. Selection and identification of Conjugation Induced Genes (COI).

Expressed sequence tags (ESTs) from vegetative (asexually growing cells and starved cells) and conjugative *Tetrahymena* cells were compared (see Materials and Method) and 58 genes that were represented only and most abundantly in the conjugative EST library (# ESTs, fourth column) where selected as conjugation-specific candidate genes. Expression of 48 of those genes was analyzed by RT-PCR, using total RNA extracted from exponentially growing (E), 16 h starved (S) or conjugating cells (from to 2h to 12 hours post-mixing). Twenty-two genes were confirmed to be conjugation-specific, since they were exclusively expressed in conjugation, with no detectable expression in vegetative stages. We named these genes “COI” (COnjugation-Induced) genes. Grey = vegetatively expressed genes. Light blue = Genes with reported function in DNA elimination (*TWI1, PDD1, CNJB, GIW1, PDD2, PDD3, EMA1* and *LIA1*) or with role in vegetative growth and meiosis (*RPB2, MND1*A gene). Pink = Conjugation Induced (*COI*) genes.

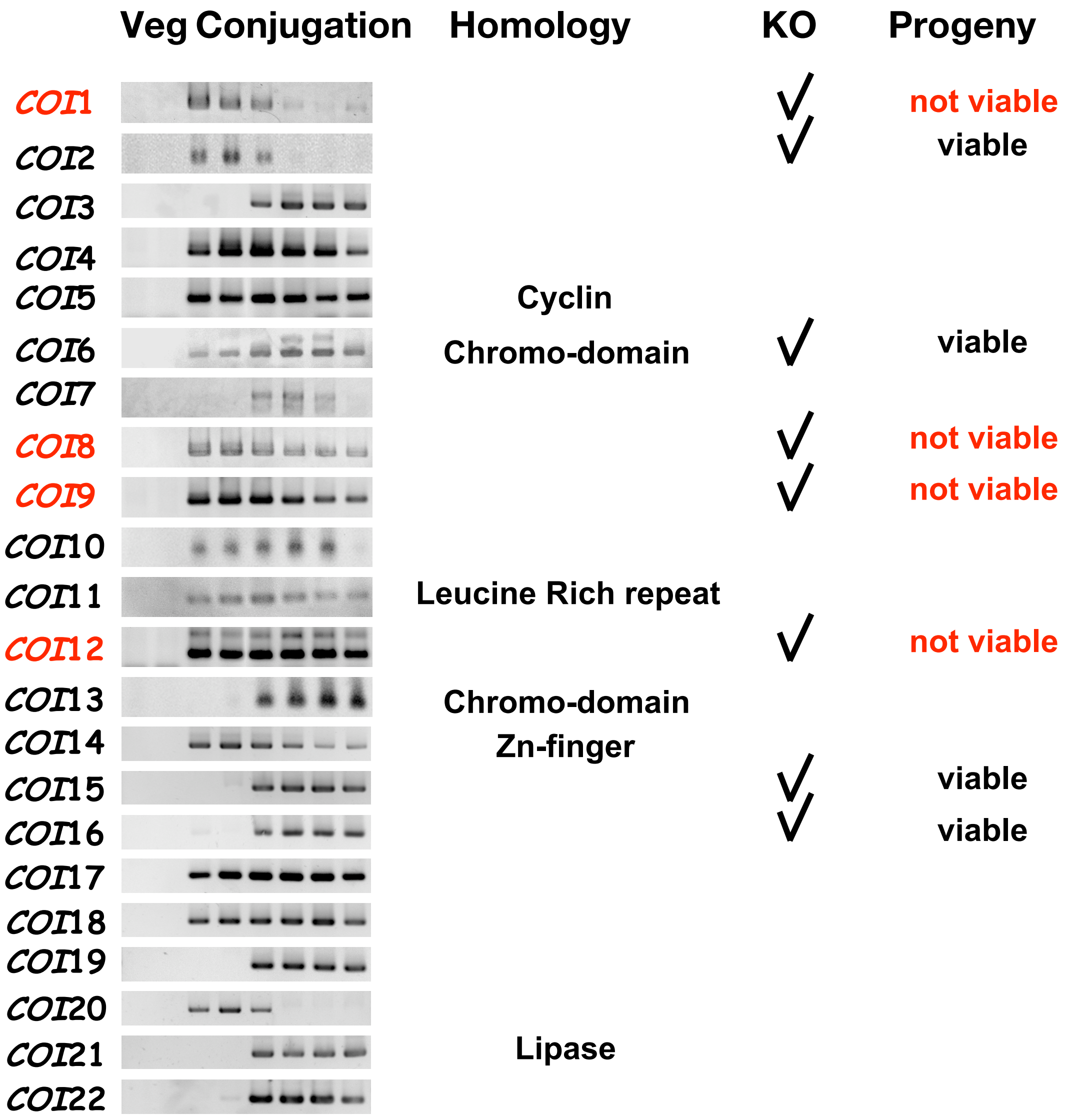
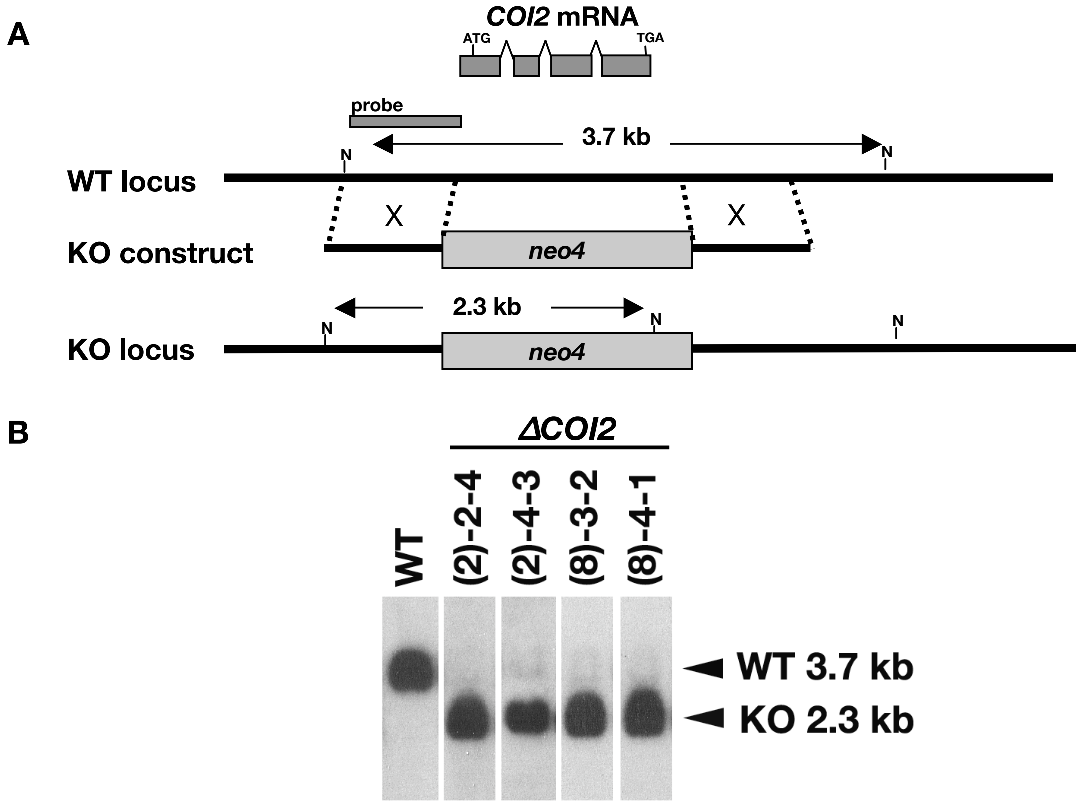


Figure 9. Gene knockout (KO) and progeny viability analysis of selected *COI* genes.

Eight *COI* genes have been so far characterized by gene knock-out (KO) and progeny viability analysis. *COI8, COI9,, COI12* produce unviable progeny after conjugation- which is in most case associated to impaired DNA elimination.

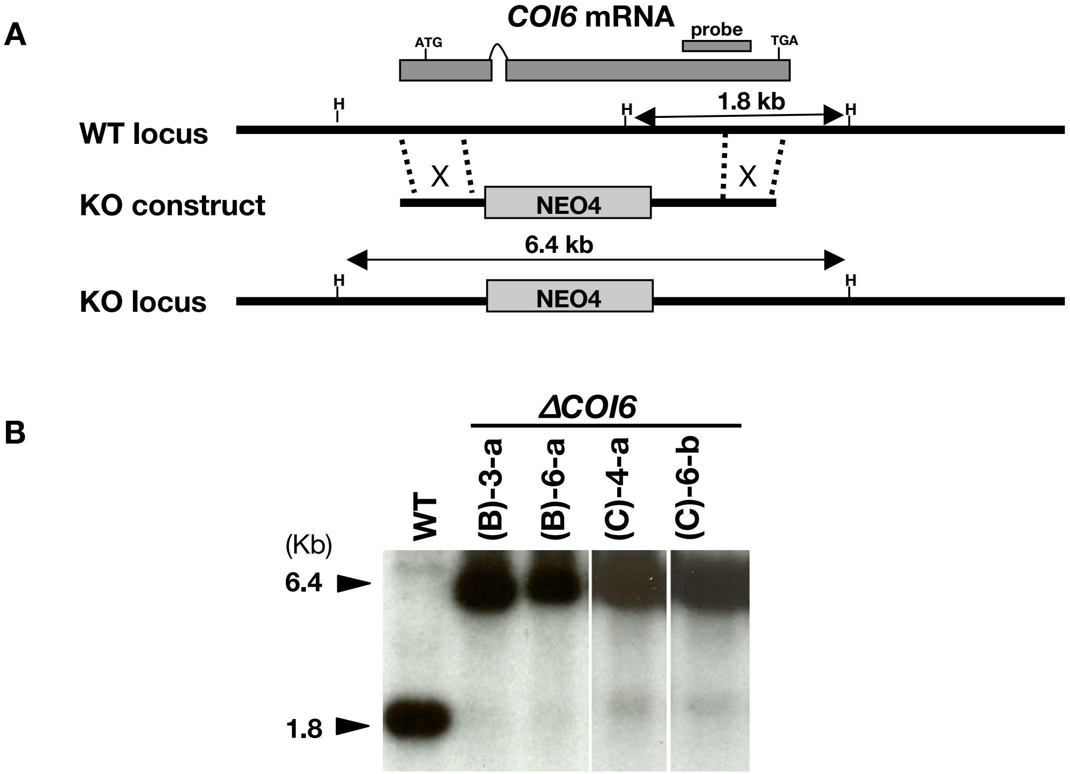


C

Type of mating cells	Percentage of cells producing viable progeny
B2086 x CU428	63.5 (n=96)
ΔCOI2 2-2-4 x *ΔCOI2 8-3-2*	57.3 (n=192)
ΔCOI6 2-2-4 x *ΔCOI6 8-4-1*	58.9 (n=192)
ΔCOI6 2-4-3 x *ΔCOI6 8-3-2*	49.5 (n=192)
ΔCOI6 2-4-3 x *ΔCOI6 8-3-2*	67.7 (n=192)

Figure 10. Construction and Characterization of *COI2* somatic KO and mutant strains

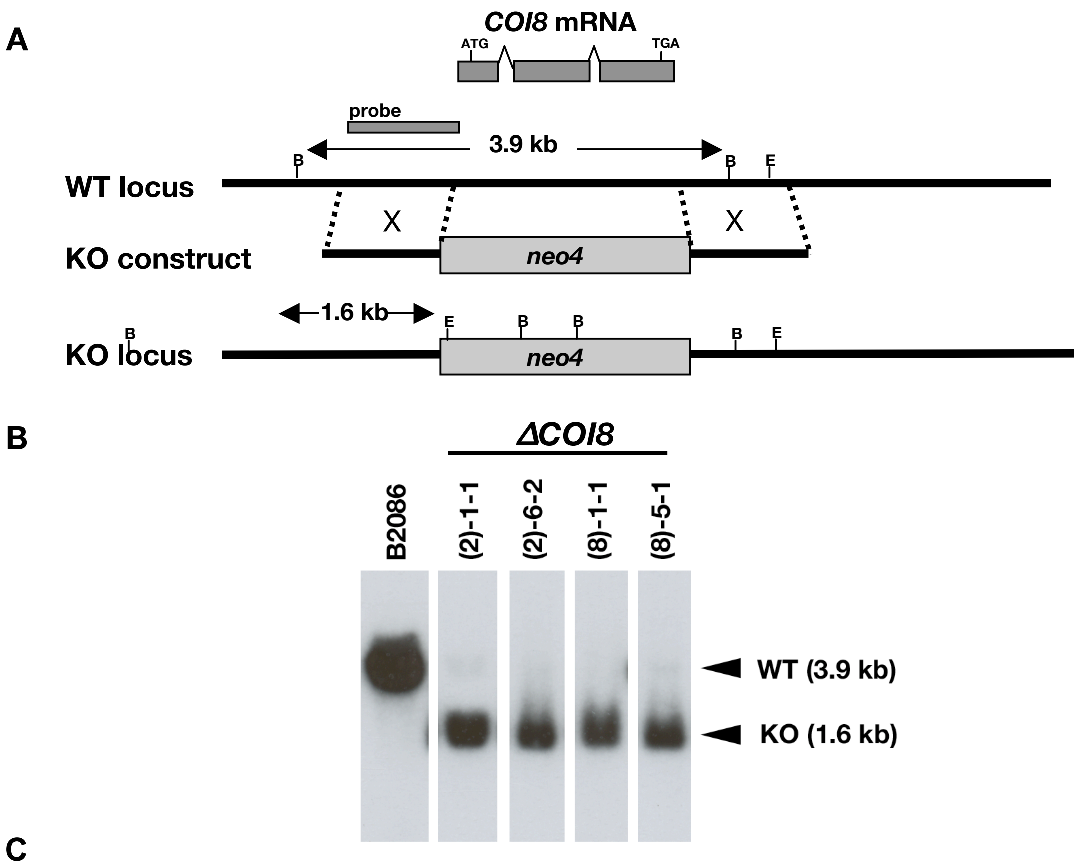
(A) Schematic drawings of the *COI2* locus and the KO construct. The entire *COI2* coding sequence was replaced by the neo4 cassette that confers paromomycin resistance in *Tetrahymena*. Upon transformation, the KO construct was introduced into the *COI2* locus by homologous recombination. (B) Southern hybridization of *COI2* KO strains.Total DNA isolated from wild-type (WT) or *COI2* KO strains (*ΔCOI2*) was digested with NcoI (N in (A)) and the blot was hybridized with the probe shown in (A). Positions of the bands for wild-type (WT) and disrupted (KO) loci are indicated with arrowheads. (C) *COI2* KO produce viable progeny. At 6~8 hr post-mixing, single mating pairs were placed into drops of SPP medium and incubated for ~60 hr at 30°C. Completion of conjugation was confirmed by testing for expression of the marker specific for newly developed Macs (6-methyl purine). Most of *COI2* KO single pairs resume vegetative growth normally.



C

Type of mating cells	Percentage of cells producing viable progeny
B2086 x CU428	70.0 (n=192)
ΔCOI6 B-3-a x *ΔCOI6 C-4-a*	69.2 (n=192)
ΔCOI6 B-3-a x *ΔCOI6 C-6-b*	76.0 (n=192)
ΔCOI6 B-6-a x *ΔCOI6 C-4-a*	73.0 (n=192)
ΔCOI6 B-6-a x *ΔCOI6 C-6-b*	66.8 (n=192)

Figure 11. Construction and Characterization of *COI6* somatic knock-out strains
(A) Schematic drawings of the *COI6* locus and the KO construct. The entire *COI6* coding sequence was replaced by the neo4 cassette that confers paromomycin resistance in *Tetrahymena*. Upon transformation, the KO construct was introduced into the *COI6* locus by homologous recombination. (B) Southern hybridization of *COI6* KO strains.Total DNA isolated from wild-type (WT) or *COI6* KO strains (Δ*COI6*) was digested with HindIII (H in (A)) and the blot was hybridized with the probe shown in (A). Positions of the bands for wild-type (WT) and disrupted (KO) loci are indicated with arrowheads. (C) *COI6* KO produce viable progeny. At 6~8 hr post-mixing, single mating pairs were placed into drops of SPP medium and incubated for ~60 hr at 30°C. Completion of conjugation was confirmed by testing for expression of the marker specific for newly developed Macs (6-methyl purine). Most of *COI6* KO single pairs resume vegetative growth normally.



C

Type of mating cells	Percentage of cells producing viable progeny
B2086 x CU428	57.8 (n=192)
ΔCOI8 2-1-1 x *ΔCOI1 8-1-1*	0 (n=480)
ΔCOI8 2-1-1 x *ΔCOI8 8-5-1*	0 (n=192)
ΔCOI8 2-6-1 x *ΔCOI8 8-1-1*	0 (n=192)
ΔCOI8 2-6-1 x *ΔCOI6 8-5-1*	0 (n=480)

Figure 12. Construction and Characterization of *COI8* somatic KO and mutant strains
(A) Schematic drawings of the *COI8* locus and the KO construct. The entire *COI8* coding sequence was replaced by the *neo4* cassette that confers paromomycin resistance in *Tetrahymena*. Upon transformation, the KO construct was introduced into the *COI8* locus by homologous recombination. (B) Southern hybridization of *COI8* KO strains.Total DNA isolated from wild-type (WT) or *COI8* KO strains (*ΔCOI8*) was digested with BglII and EcoRI (B and E in (A)) and the blot was hybridized with the probe shown in (A). Positions of the bands for wild-type (WT) and disrupted (KO) loci are indicated with arrowheads. (C) *COI8* KO produce no viable progeny. At 6~8 hr post-mixing, single mating pairs were placed into drops of SPP medium and incubated for ~60 hr at 30°C. Completion of conjugation was analyzed by testing for expression of the marker specific for newly developed Macs (6-methyl purine).

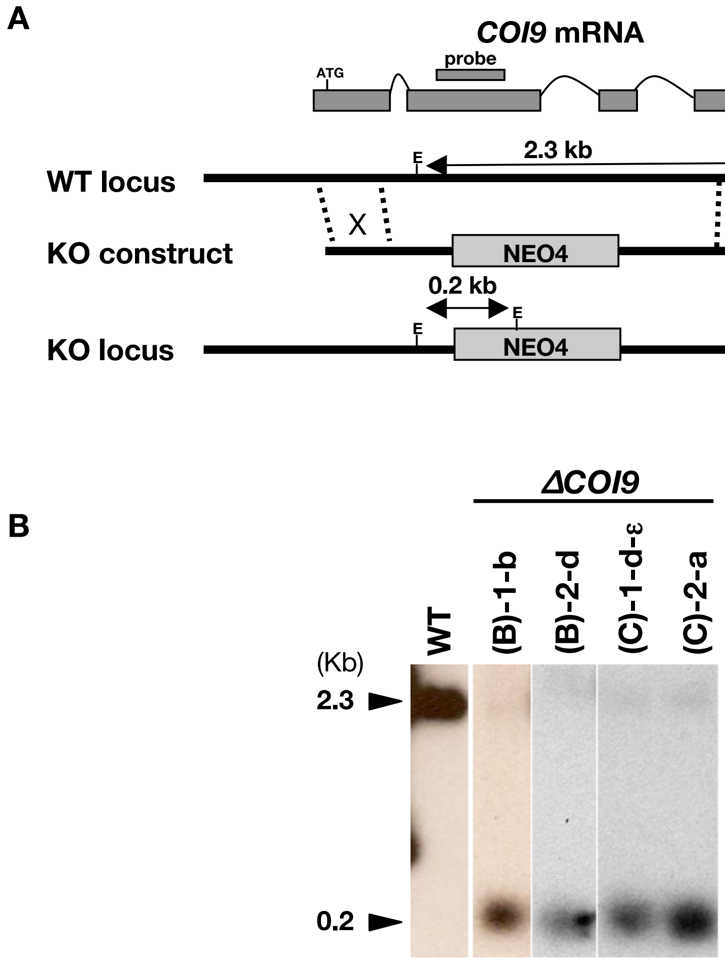


Figure 13. Construction and characterization of *COI9* somatic knock-out strains
(A) Schematic drawings of the *COI9* locus and the KO construct. The entire *COI9* coding sequence was replaced by the neo4 cassette that confers paromomycin resistance in Tetrahymena. Upon transformation, the KO construct was introduced into the *COI9* locus by homologous recombination. (B) Southern hybridization of *COI9* KO strains.Total DNA isolated from wild-type (WT) or *COI9* KO strains (ΔCOI9) was digested with EcoR1 (E in (A)) and the blot was hybridized with the probe shown in (A). Positions of the bands for wild-type (WT) and disrupted (KO) loci are indicated with arrowheads.

A

Type of mating cells	Percentage of cells producing viable progeny
B2086 x CU428	**77 .6 (n=192)**
ΔCOI9 B-1-b x ΔCOI9 C-1-d-ε	**5.1 (n=192)**
ΔCOI9 B-1-b x ΔCOI9 C-2-a	**4.7 (n=192)**
ΔCOI9 B-2-d x ΔCOI9 C-1-d-ε	**3.8 (n=192)**
ΔCOI9 B-2-d x ΔCOI9 C-2-a	**4.2 (n=192)**

B

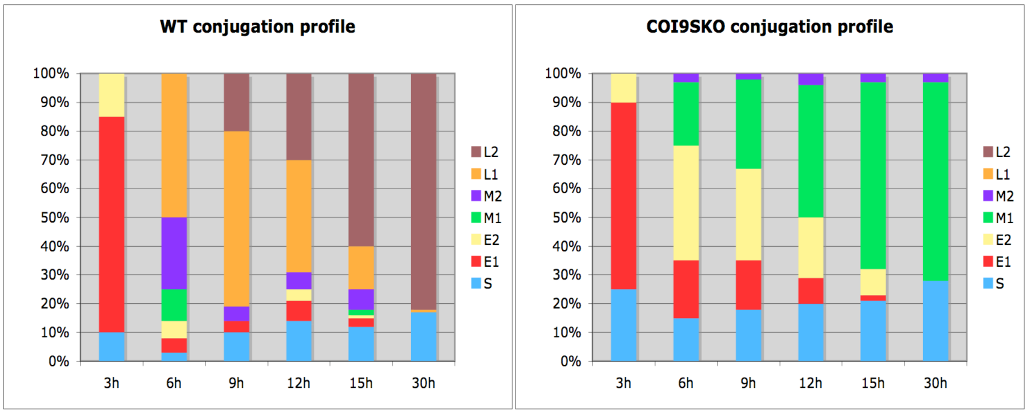


Figure 14. *COI9* KO strain fail to complete conjugation and produce viable progeny.
A) *COI9* KO cells show reduced progeny viability. At 6~8 hr post-mixing, single mating pairs were placed into drops of SPP medium and incubated for ~60 hr at 30°C. Completion of conjugation was confirmed by testing for expression of the marker specific for newly developed Macs (6-methyl purine). Most single pairs of *COI9* KO cells fail to resume vegetative growth and eventually die. *B)* Developmental profiles of conjugation in WT and *COI9* KO strains. Starved wild type cells (WT, B0286 X CU428) and *COI9* KO strains (Δ*COI9*) were mixed and stages of conjugations were observed by DAPI staining. The stages I categorized were: S, single cells (unmated); E1, Pre-Meiosis; E2, Meiosis; M1,Prezygotic; M2, Postzygotic; L1, Mac Development; L2, Pair separation (2Mics); L3, Mic Elimination. See Fig. 2 for the developmental stages. Disruption of *COI9* causes arrest during or right after meiosis (E2-M1 stage).

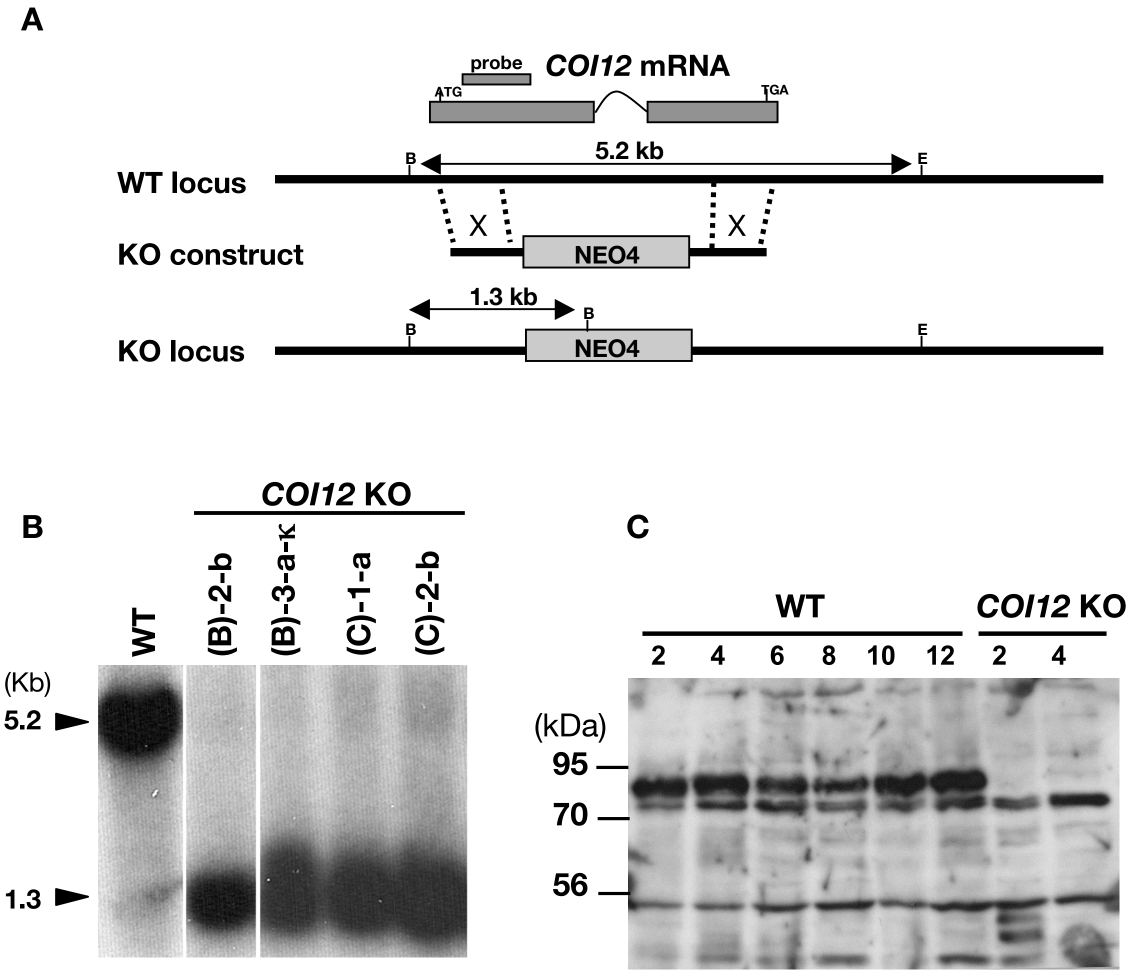


Figure 15. Construction and Characterization of *COI12* somatic KO and mutant strains

(A) Schematic drawings of the *COI12* locus and the KO construct. The entire *COI12* coding sequence was replaced by the neo3 cassette that confers paromomycin resistance in Tetrahymena. Upon transformation, the KO construct was introduced into the *COI12* locus by homologous recombination. (B) Southern hybridization of *COI12* KO strains.Total DNA isolated from wild-type (WT) or COI12 KO strains [ΔCOI12: (B)-2-b, (B)-3-a-κ, (C)-1-a, (C)-2-b)] was digested with BglII and EcoRV (B and E in (A)) and the blot was hybridized with the probe shown in (A). Positions of the bands for wild-type (WT) and disrupted (KO) loci are indicated with arrowheads. (C) Expression of Coi12p in WT and *COI12* KO strains. Coi12p expression in mating wild-type (WT; 2-12 hr post-mixing) and *COI12* KO strains (*COI12* KO; 2, 4 hr post-mixing) was analyzed by Western blot using anti-Coi12p antiserum. No COI12p was detected in *COI12* KO strains confirming the complete replacement of *COI12* loci with the KO construct.

A

Type of mating cells	Percentage of cells producing viable progeny
B2086 x CU428	**72.0 (n=192)**
ΔCOI12 B-2-b* x *ΔCOI12 C-1-a	**0 (n=192)**
ΔCOI12 B-2-b* x *ΔCOI12 C-3-b	**0 (n=192)**
ΔCOI12 B-3-a-κ* x *ΔCOI12 C-1-a	**0 (n=192)**
ΔCOI12 B-3-a-κ* x *ΔCOI12 C-1-a	**0 (n=192)**

B

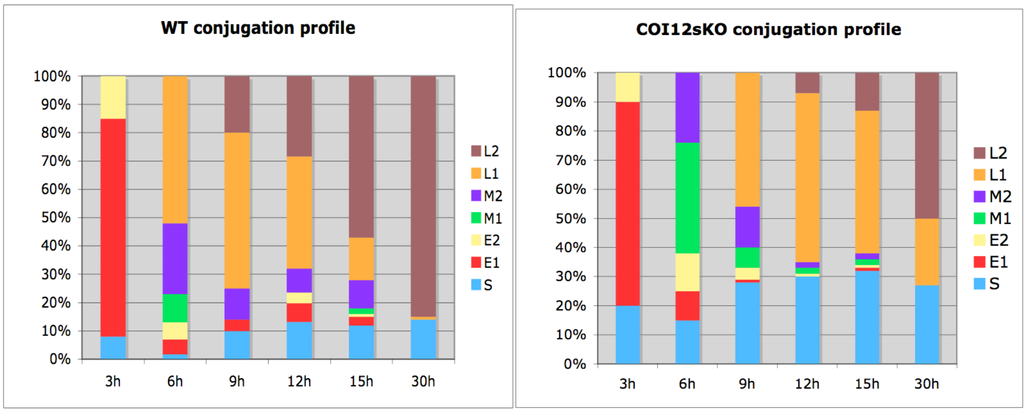


Figure 16. *COI12* KO strain fail to complete conjugation and produce viable progeny. A) *COI12* KO cells fail to produce viable progeny. At 6~8 hr post-mixing, single mating pairs were placed into drops of SPP medium and incubated for ~60 hr at 30°C. Completion of conjugation was confirmed by testing for expression of the marker specific for newly developed Macs (6-methyl purine). All single pairs of *COI12* KO cells fail to resume vegetative growth and eventually die. *B)* Developmental profiles of conjugation in WT and *COI12* KO strains. Starved wild type cells (WT, B0286 X CU428) and *COI12* KO strains (Δ*COI12*) were mixed and stages of conjugations were observed by DAPI staining. The stages I categorized were: S, single cells (unmated); E1, Pre-Meiosis; E2, Meiosis; M1,Prezygotic; M2, Postzygotic; L1, Mac Development; L2, Pair separation (2Mics); L3, Mic Elimination. See Fig. 2 for the developmental stages. Disruption of *COI12* causes arrest at Pair separation (L2) stage.

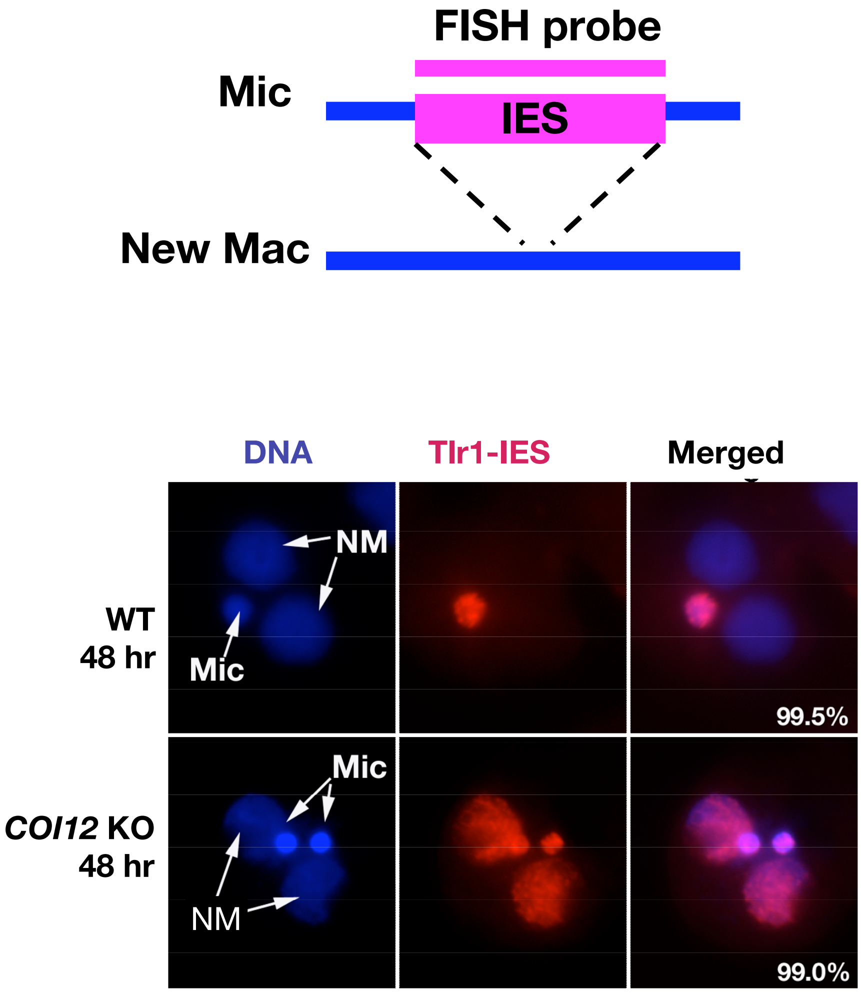


Figure 17. DNA elimination is inhibited in the progeny of *COI12* KO cells
Exconjugants (sexual progeny) of wild-type (WT) and *COI12* KO strains at 48 hr post mixing were used to detect Tlr1-IES elements by fluorescent in situ hybridization (middle, red). DNA was stained with DAPI (left, blue). 99.5% of exconjugants from wild-type cells showed Tlr1-FISH staining only in the micronucleus. In contrast, 99.0% of exconjugants from *COI12* KO strains showed Tlr1-FISH staining in both the micronuclei and the newly developed macronuclei, indicating failure to eliminate this IES. The micronuclei (Mic) and the newly developed macronuclei (NM) are marked with arrowheads.

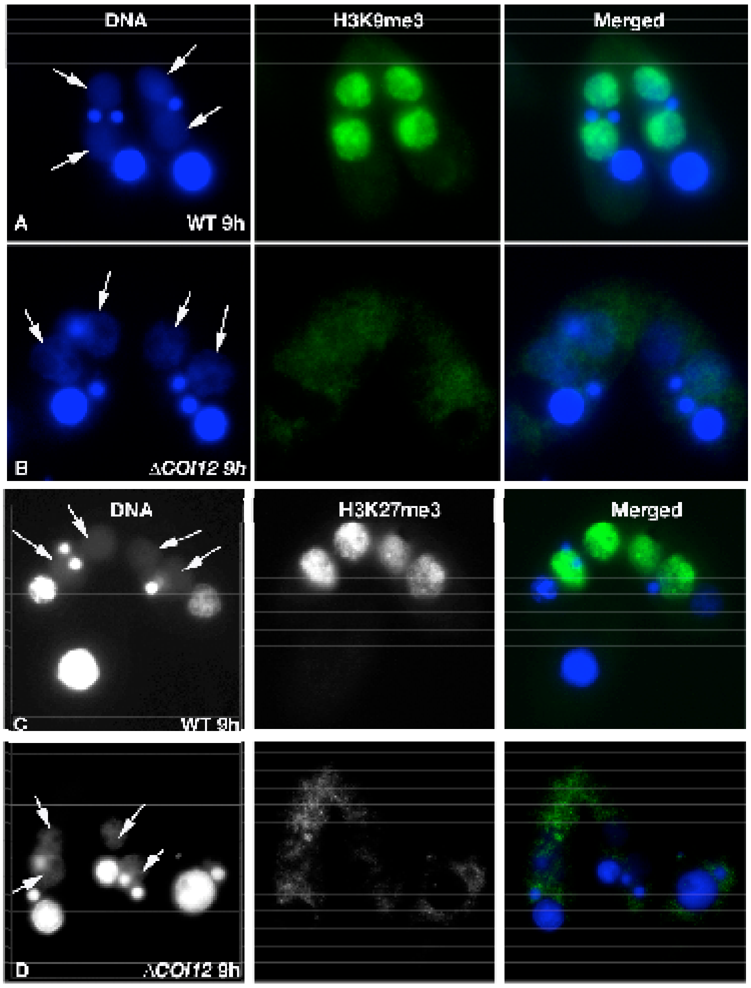


Figure 18. H3K9 and H3K27 methylation in new developing Macs are impaired in *COI12* KO strains
Wild type (WT) and *COI12 (COI12)* cells in conjugation at 9 hr post-mixing were processed for indirect immunofluorescent staining (middle, green) using anti H3K9me 3(A-B) and anti H3K27me 3 (D-E) antibodies. DNA was stained by DAPI (left, blue). Arrowheads indicate developing new Macs.

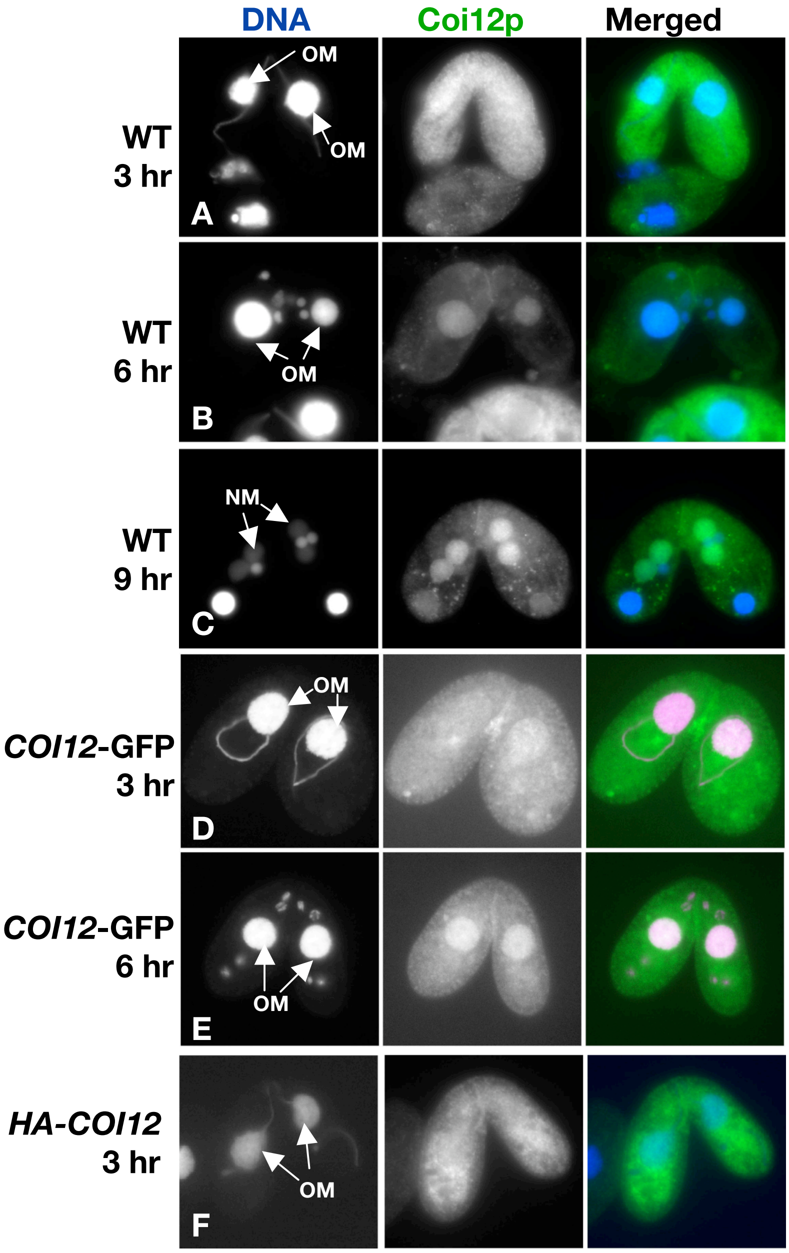


Figure 19. Coi12p localizes in cytoplasm and macronuclei.
WT strains were fixed at 3 hr, 6 hr and 9 hr post-mixing and analyzed by immunostaining using anti-Coi12 antibody (A, B, C). In WT strains, Coi12p localizes in cytoplasm at early stages of conjugation (3h). Coi12p is detected mainly in the parental Mac at mid conjugation (6 hr) and in new Macs at late stages (9 hr). The same localization pattern of Coi12p was observed using *COI12*GFP (D-E) or *COI12HA* tagged strains and anti-HA antibodies (F). However, in both *COI12-GFP* and HA strains, no fluorescent signal is detected at late stages of conjugation. OM= old Mac; NM= new mac.

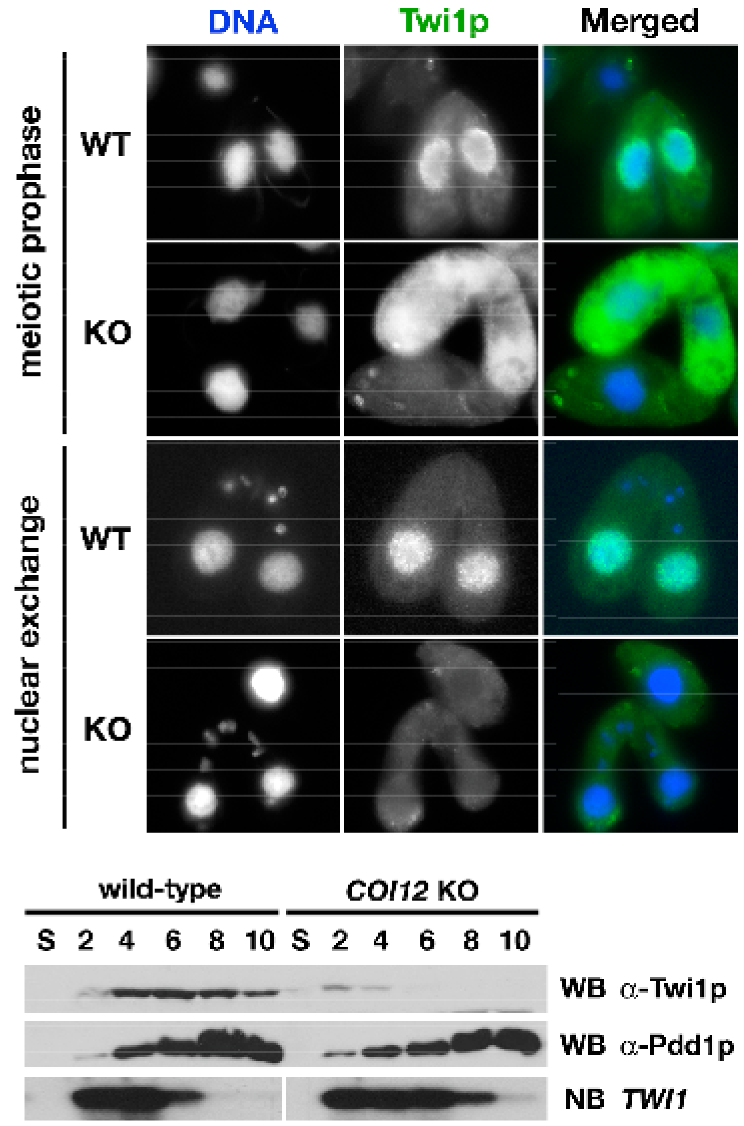


Figure 20. *COI12* is required for both macronuclear localization and stable accumulation of Twi1p
(A-D) Wild-type (WT in A and C) and *COI12* KO (KO, in B and D) cells in conjugation at the indicated stages were processed for indirect immunofluorescent staining. Twi1p was localized using anti-Twi1p antiserum (green, middle) and DNA was stained by DAPI (blue, right). Macronuclei are marked by arrowheads. (E) Accumulation of Twi1p in starved (S) and conjugating (from 2 to 10 hours post-mixing) wild-type (WT) and *COI12* KO strains was analyzed by Western blot. As control, accumulation of Pdd1p (lower panel), which is also required for DNA elimination, was analysed. On the bottom, expression of TWI1 mRNA was analysed by Northern blot.

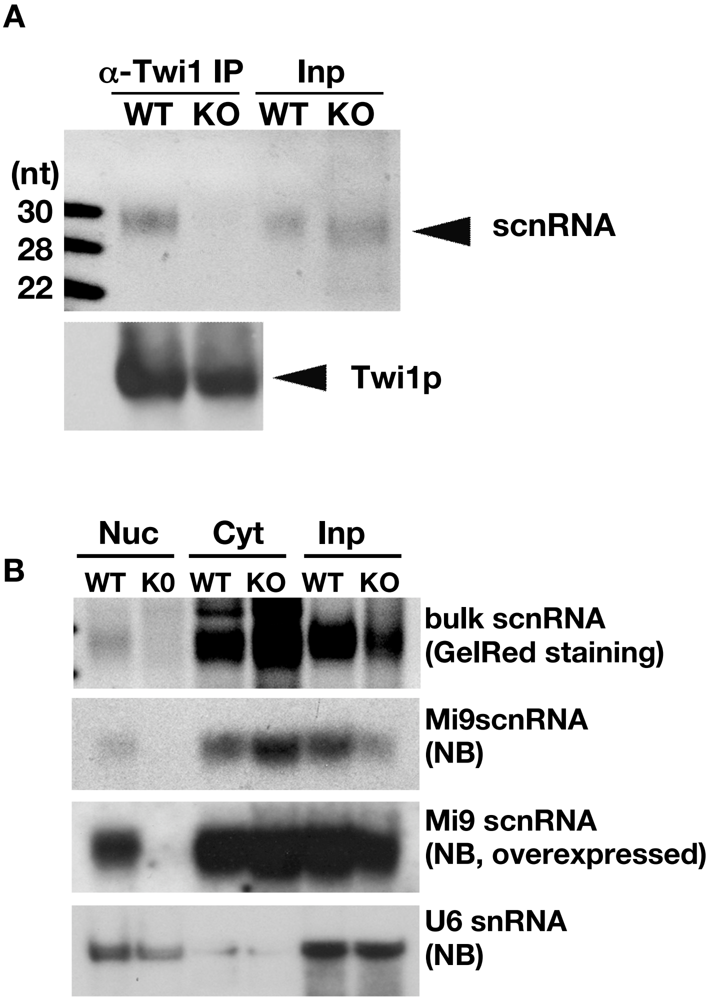


Fig. 21. ScnRNA are normally exported from the micronucleus into the cytoplasm but they are not loaded into Twi1p in *COI12 KO strains.*
(A) ScnRNA are not loaded into Twi1p in *COI12* KO strains. Twi1p-containing complexes from wild-type (WT) and *COI12* knockout (KO) strains at 2.5 hr post-mixing were immunoprecipitated using an anti-Twi1p antibody and co-immunoprecipitated RNA was analyzed by denaturing gel electrophoresis, followed by staining with GelRed. No scnRNA was detected in complex with Twi1p in *COI12* KO strains. Amount of total scnRNA (input, right) and immunoprecipitated Twi1p (bottom) was comparable between WT and Δ*COI12* strains. ScnRNAs are exported normally from the micronucleus to the cytoplasm in *COI12* KO strains. Nuclear (Nuc) and cytoplasmic (Cyt) RNA were fractionated in conjugating wild-type (WT) and *COI12* KO (KO) cells at 3 hr post-mixing. Fractionated RNA was analysed by denaturing gel electrophoresis, followed by staining with GelRed (top) or Northern blot using Mi9 scnRNA as a probe (middle). As a marker for the purity of the separation, U3 snRNA (bottom) was analysed in both nuclear and cytoplasmic fractions.

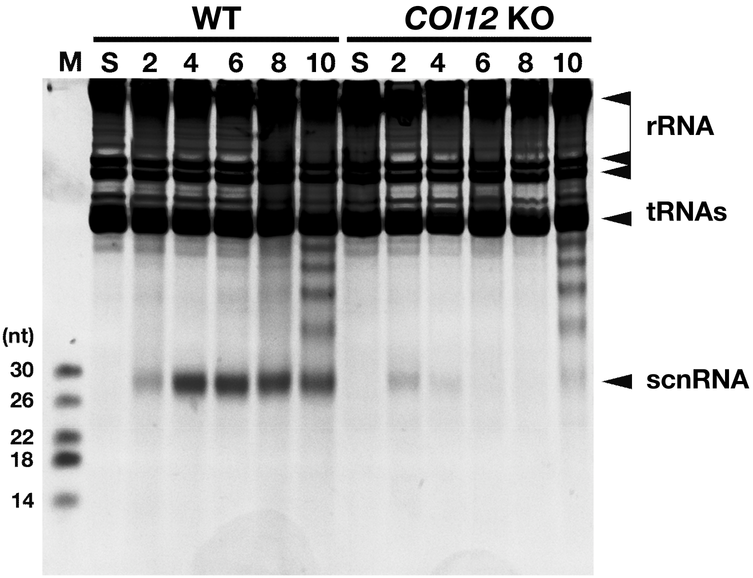


Figure 22. Coi12p is required for stable accumulation of scnRNA.

Total RNA was extracted from starved (S) or conjugating (2, 4, 6, 8, and 10 hr post-mixing) wild-type (WT) or *COI12* KO strains and separated in sequencing gels. Bulk scnRNA (~28 nt) was visualized by staining with GelRed.

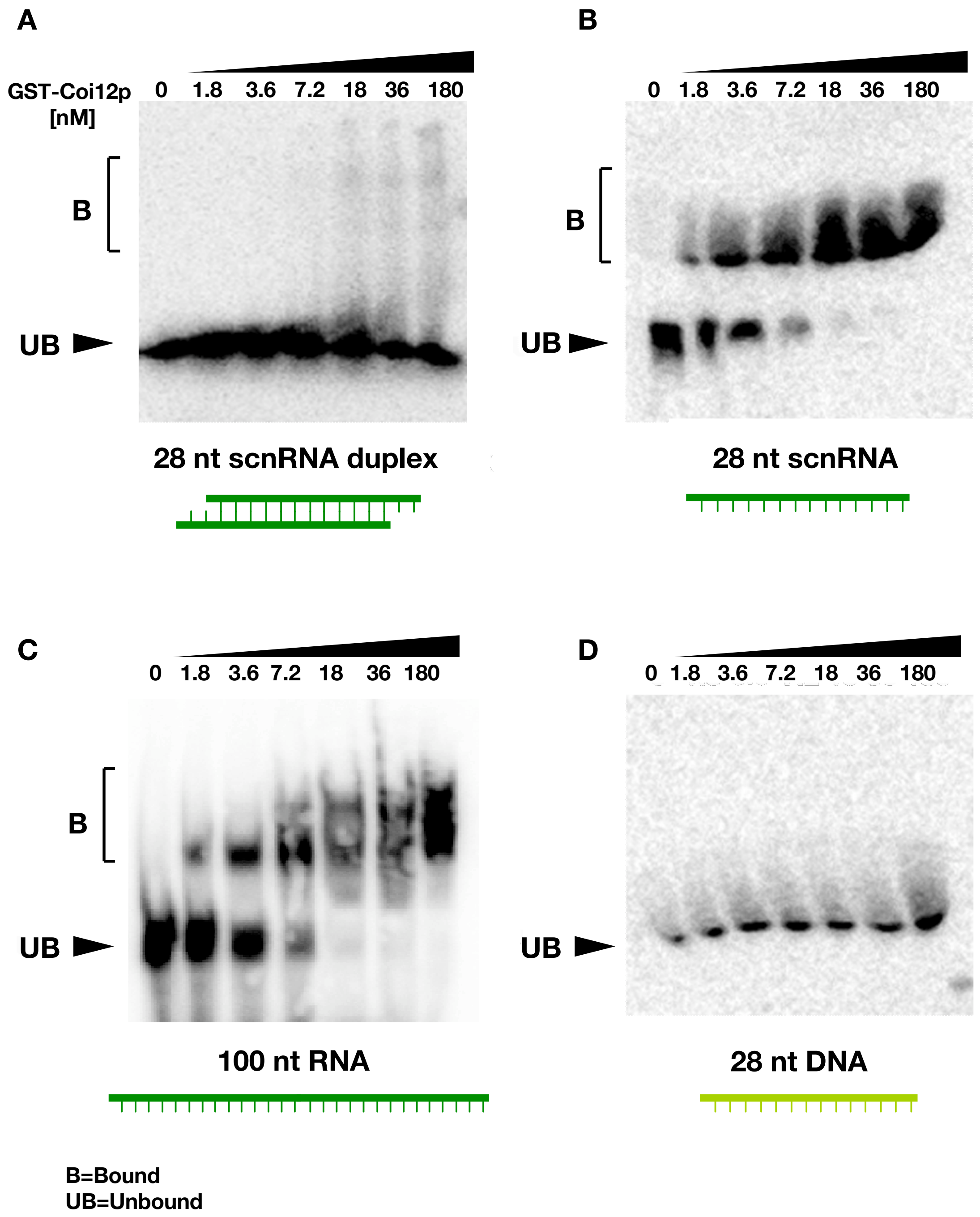


Figure 23. Coi12p is a single-stranded RNA binding protein.
Recombinant Coi12p at the indicated, increasing concentrations was incubated with ^{32}P-labeled 29 nucleotides (nt) double-stranded (ds) scnRNAs (A), 29 nt single-stranded (ss) RNAs (B), 100 nt ssRNAs (C) or 29 nt ssDNA (D) and formation of ribonucleoprotein complex was analysed by electrophoretic mobility shift assay. Bound (B) and unbound (UB) RNA is indicated with brackets and arrowheads respectively.

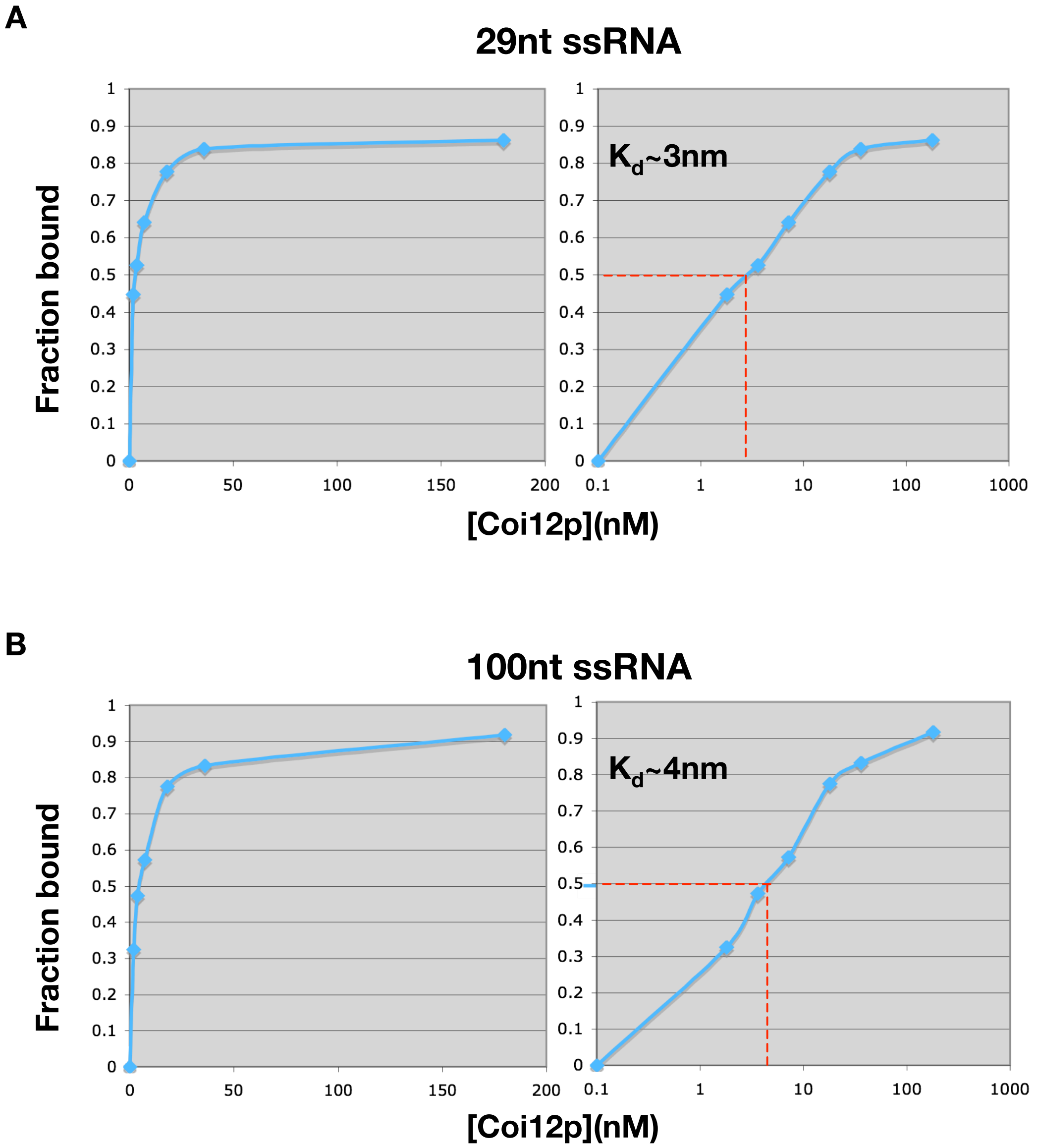


Figure 24. Coi12p binds with high, comparable affinity to 28 nt and 100 nt long ssRNA .

Hyperbolic (left) and logaritmic (right) curves defining the bimolecular binding between Coi12p and 29 nt ssRNA (A) or between Coi12p and 100 nt ssRNA (B). The dashed lines indicate that the dissociation constant (Kd) equals the concentration of Coi12p at which the fraction of ssRNA bound is 0.5.

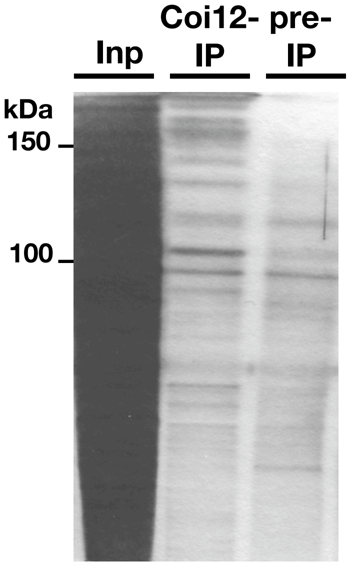
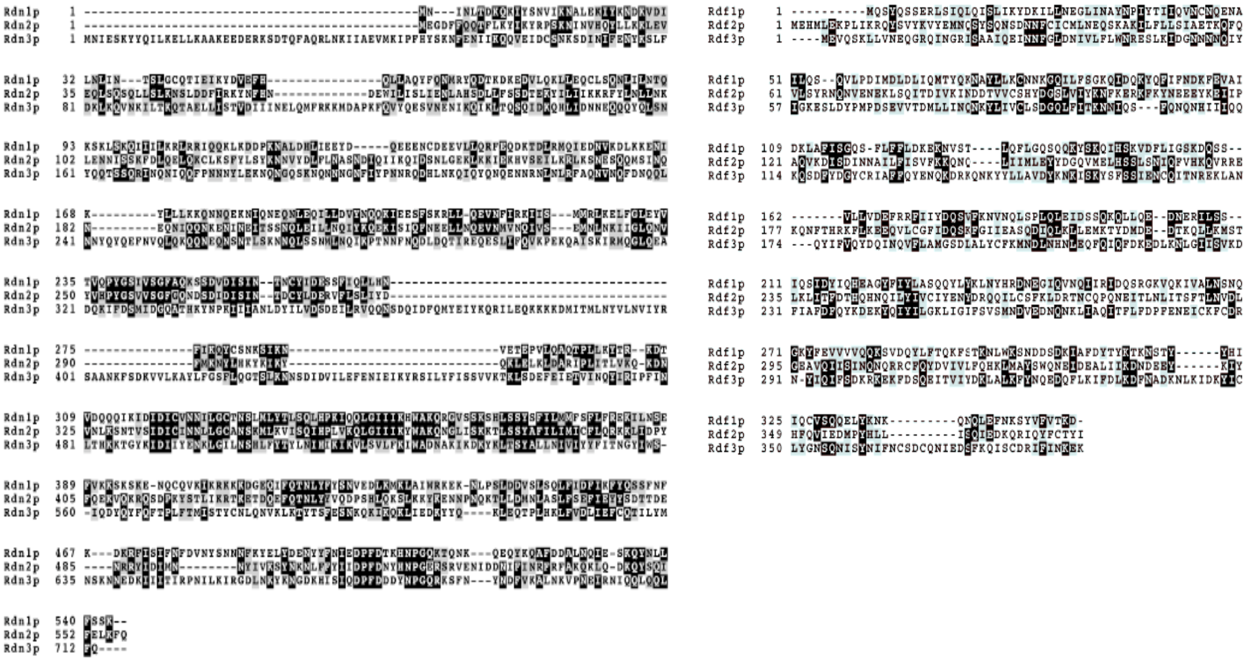


Figure 25. Co-immunoprecipitation of Coi12p-associated proteins.
Coi12p was immunoprecipitated using antibody anti-Coi12p from WT conjugating cells harvested at 3 hr post-mixing. The purified protein complex was separated by SDS PAGE, visualized by silver staining and analyzed by mass spectrometry (Appendix 1).

A

Identified Proteins	homology/domains	Accession #	MW	# unique peptides	MW/pep	Coverage
TTHERM_00402050, Coi12p		gi|229595943	64	14	4.6	33%
TTHERM_01207560, Rdf3p	RdRP-associated factor	gi|118377447	87	6	14.5	9%
TTHERM_01394310, Rdn3p	Nucleotidyltransferase	gi|118394510	85	2	42.5	2%
TTHERM_01216140, Dcr1p	Dicer	gi|118352386	247	3	82.3	2%

B



C

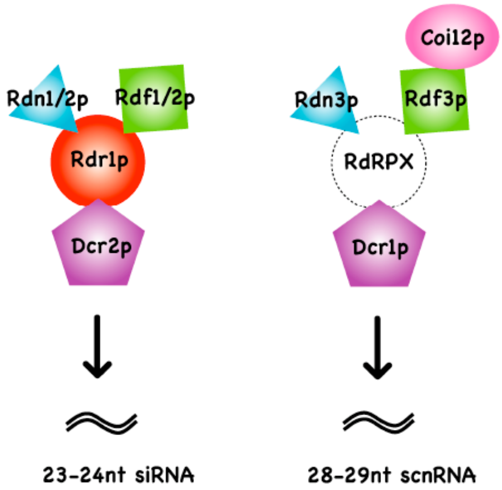


Figure 26. Coi12p-associated complex may be a novel RNA-dependent RNA polymerase complex with function in DNA elimination

A) Proteins identified in Coi12p immunocomplex with homology to RdRPP factors. B) Sequence alignment of Coi12-p associated Rdn3 (left) and Rdf3 (right) with Rdf1-2 and Rdn1-2 respectively. MW= molecular weight. C) Model of composition and function of vegetative and conjugative RNA-dependent RNA polymerase complexes I (RDRC) in *Tetrahymena*. Coi12p-associated RNA-dependent RNA polymerase (RdRPx) has not been yet identified.

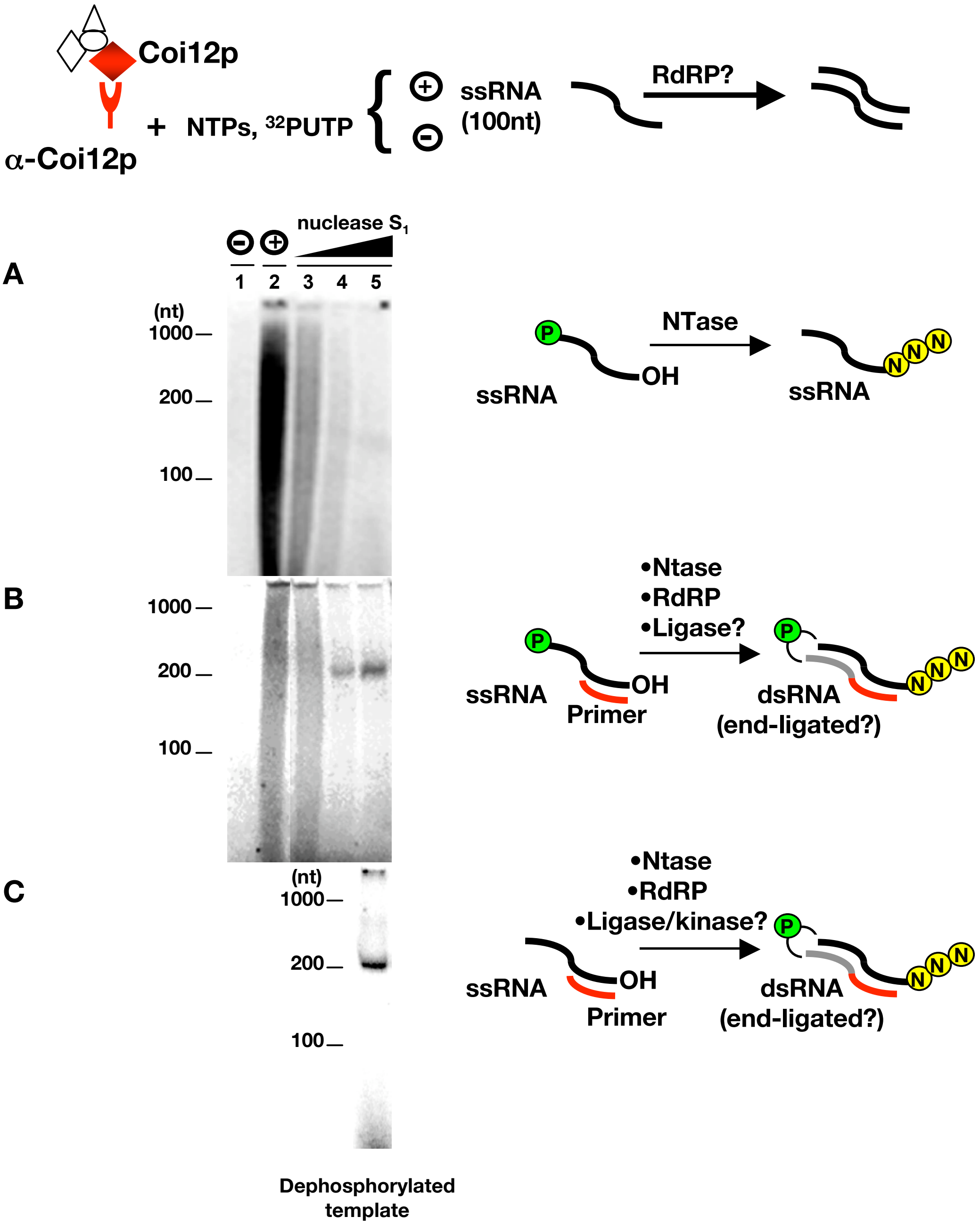


Figure 27. A primer-dependent RdRP activity is associated with Coi12p.
Coi12p-coimmunoprecipitated proteins were incubated with P^{32}UTP and cold NTPs in presence (lane 2, +) or absence (lane 1, -) of 100 ssRNA template alone (A) or with a 29 nt primer at annealed at its 3’end (B). Reaction products were analyzed by polyacrilamide gel electrophoresis and by autoradiography, before or after digestion with increasing concentrations of nuclease S1 (S1, lanes 3-5)- which is specific for single-stranded nucleic acids. The same reaction was performed using dephosphorylated, primer-annealed template (C) to investingate the involvement of a RNA-ligase in the reaction.

Species	Argonaute-Piwi	Dicer-like	RdRP	Role in RNAi
Excavata				
Giardia intestinalis	+	+	+	?
Trypanosoma brucei	+	–	–	+
Trypanosoma cruzi	–	–	–	NA
Leishmania major	–	–	–	NA
Chromalveolata				
Paramecium tetraurelia	+	+	+	?
Tetrahymena thermophila	+	+	+	+
Plasmodium falciparum	–	–	–	NA
Phytophthora sojae	+	+	+	?
Thalassiosira pseudonana	+	–	–	?
Rhizaria				
Data not available				
Archaeplastida				
Cyanidioschyzon merolae	–	–	–	NA
Chlamydomonas reinhardtii	+	+	–	+
Arabidopsis thaliana	+	+	+	+
Oryza sativa (japonica)	+	+	+	+
Amoebozoa				
Dictyostelium discoideum	+	+	+	+
Entamoeba histolytica	+	–	+	?
Opisthokonta				
Saccharomyces cerevisiae	–	–	–	NA
Schizosaccharomyces pombe	+	+	+	+
Neurospora crassa	+	+	+	+
Aspergillus nidulans	+	+	+	?
Caenorhabditis elegans	+	+	+	+
Drosophila melanogaster	+	+	–	+
Anopheles gambiae	+	+	–	+
Strongylocentrotus purpuratus	+	+	–	?
Ciona intestinalis	+	+	–	?
Homo sapiens	+	+	–	+

Figure 28. Distribution of RNA-dependent RNA polymerases proteins and activities among eukaryotes.
Expression and role in RNAi of Argonaute, Dicer, and RdRP proteins is indicated for each species listed on the right and grouped into 6 eukaryotic families (in light green). Role in RNAi is inferred from the phenotype of mutant or RNAi knock-down strains defective in at least one of the RNAi effectors. +, presence; -, absence; ?, lack of evidence; NA, not applicable.

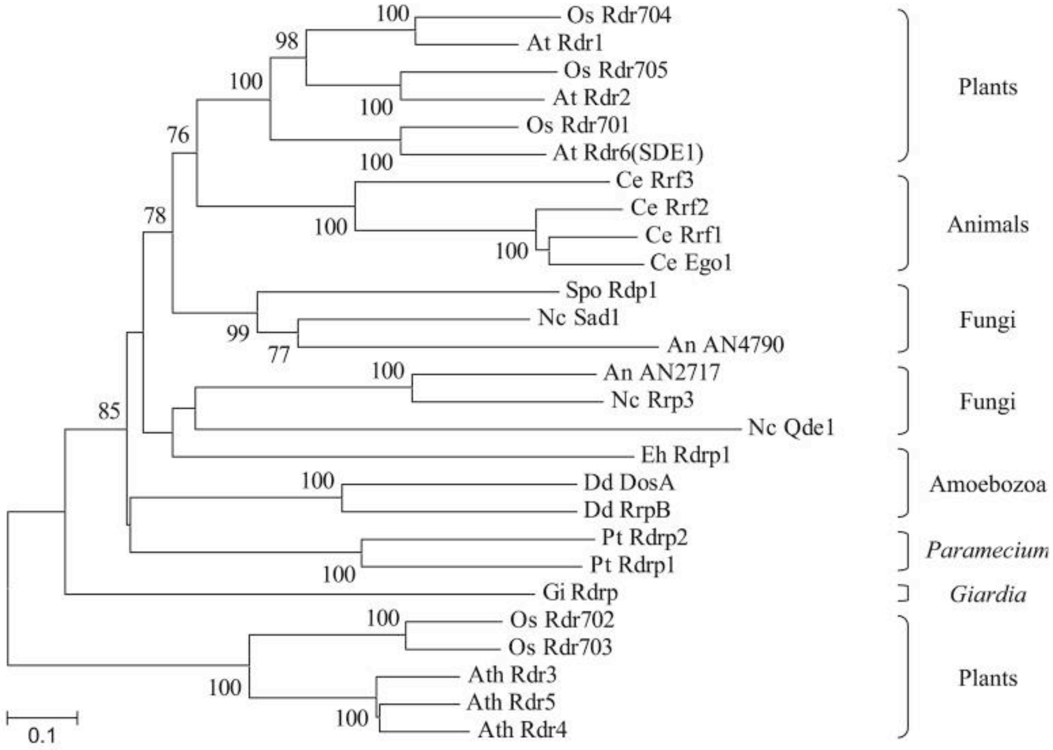


Figure 29. Phylogenetic tree of eukaryotic RNA-dependent RNA polymerases
An alignment of sequences corresponding to the RdRP domain was used to construct a Neighbor-Joining tree and its robustness was assessed by a bootstrap test based on 1,000 pseudoreplicates. Bootstrap values higher than 60% are shown on the nodes of the tree. Species are indicated by a two-letter abbreviation preceding the name of each protein. At *A. thaliana*; Ce *C. elegans*; Dd *D. discoideum*; Eh *E. histolytica*; Gi *G. intestinalis*; Nc *N. crassa*; Os *O. sativa*; Pt *P. tetraurelia*; Spo *S. pombe*.

Appendix 1

gray: related to rRNA synthesis, transcription, splicing
pink: RdRP-related

Protein ID	homology/domains/other names	Accession #	MW	# peptides	MW/pep
TTHERM_01289080	Sm protein	gi|118394903	13	5	2.6
TTHERM_00557860	coiled-coil domain	gi|229594708	55	16	3.4
TTHERM_00820690	-	gi|118365916	21	6	3.5
TTHERM_00585260	ATP synthase beta chain (mitochondrial?)	gi|146185860	53	14	3.8
TTHERM_01129660	Ribosomal protein L36e	gi|118387091	12	3	4
TTHERM_00290940	-	gi|118371022	45	11	4.1
TTHERM_00145490	-	gi|146163305	13	3	4.3
TTHERM_00954040	-	gi|118369648	22	5	4.4
TTHERM_00402050	Coi12p	gi|229595943	64	14	4.6
TTHERM_00149630	ribosomal protein L34	gi|118377144	14	3	4.7
TTHERM_00812780	ribosomal protein L38e	gi|146162057	10	2	5
TTHERM_00594220	ribosome biogenesis regulatory protein (RRS1)	gi|118356857	36	7	5.1
TTHERM_00285510	-	gi|118370730	52	10	5.2
TTHERM_00444510	ribosomal protein L19e	gi|146181731	21	4	5.3
TTHERM_00939210	-	gi|118397525	21	4	5.3
TTHERM_00317230	-	gi|118376444	16	3	5.3
TTHERM_00494550	ATPase or ATP binding	gi|229594539	16	3	5.3
TTHERM_00497880	ribosomal protein L3	gi|118389862	44	8	5.5
TTHERM_00189170	Histone H4, minor, putative	gi|118366755	11	2	5.5
TTHERM_00727690	RRM, RNA binding protein	gi|118378971	45	8	5.6
TTHERM_00630500	Calmodulin, EF hand family protein	gi|146181449	17	3	5.7
TTHERM_01289110	ribosomal protein L15	gi|118394925	24	4	6
TTHERM_00266529	highly similar to TTHERM_00266520	gi|146165988	18	3	6
TTHERM_00392980	-	gi|146163587	12	2	6
TTHERM_00564540	-	gi|118377751	85	14	6.1
TTHERM_00706330	-	gi|146163188	26	4	6.5
TTHERM_00248430	RRM, nucleolar phosphoprotein?	gi|118381422	33	5	6.6
TTHERM_00803640	Protein kinase domain containing protein	gi|118358332	68	10	6.8
TTHERM_00686150	Ribosomal protein S27a	gi|118384128	21	3	7
TTHERM_00929590	H/ACA small nucleolar RNP component Gar1 (RNA binding)	gi|118385169	22	3	7.3
TTHERM_00456870	Zinc finger, C2H2 type family protein	gi|118382069	22	3	7.3
TTHERM_00522820	Coi4p	gi|146165131	54	7	7.7
TTHERM_01005320	60s Acidic ribosomal protein	gi|118354178	31	4	7.8
TTHERM_00660180	HMG box family protein, Ribosomal Transcription Factor Ubtf	gi|229595100	16	2	8
TTHERM_00773800	-	gi|118398610	24	3	8
TTHERM_00918460	-	gi|229593576	16	2	8
TTHERM_00579340	-	gi|118379538	41	5	8.2
TTHERM_00444590	rRNA-processing protein, Fcf2	gi|118380316	25	3	8.3
TTHERM_00549610	RNA polymerase Rpb8 family protein	gi|118345758	17	2	8.5
TTHERM_00077420	c-myc binding protein	gi|118364924	17	2	8.5
TTHERM_00622700	-	gi|118378642	17	2	8.5
TTHERM_00325650	Ribosomal protein L7Ae	gi|118386619	18	2	9
TTHERM_00322790	-	gi|118359481	19	2	9.5
TTHERM_00551150	-	gi|229596452	77	8	9.6
TTHERM_00534010	zinc finger protein	gi|229594299	29	3	9.7
TTHERM_00942990	SKIP/SNW domain, spliceosomal component & transcriptional coregulator	gi|118397414	50	5	10
TTHERM_00285620	nascent polypeptide associated complex (NAC) domain	gi|146172717	21	2	10.5
TTHERM_00655440	carbohydrate-binding module family 20 (CBM20)	gi|229594488	32	3	10.7

Appendix 1

gray: related to rRNA synthesis, transcription, splicing
pink: RdRP-related

Identified Proteins	homology/domains/other names	Accession #	MW	# peptides	MW/pep
TTHERM_00941510	Aurora kinase	gi|118387614	155	14	11.1
TTHERM_00563860	-	gi|229594733	24	2	12
TTHERM_00780700	translation initiation factor (eIF2B/eIF5 family)	gi|146183450	24	2	12
TTHERM_00035630	Brix domain containing protein, ribosomal RNA processing?	gi|118346335	37	3	12.3
TTHERM_00327090	Zinc finger?	gi|118386709	25	2	12.5
TTHERM_00932010	-	gi|118377354	135	10	13.5
TTHERM_01212840	-	gi|118387822	27	2	13.5
TTHERM_00825440	TOP2, DNA gyrase/topoisomerase	gi|118398165	164	12	13.7
TTHERM_00938920	-	gi|118397469	833	59	14.1
TTHERM_00637800	Utp14, U3 small nucleolar RNA-associated protein	gi|118400236	86	6	14.3
TTHERM_00780490	-	gi|118386169	216	15	14.4
TTHERM_01207560	Rdf3p, (two ORFs, next to Rdf1p)	gi|118377447	87	6	14.5
TTHERM_00497330	-	gi|146184163	44	3	14.7
TTHERM_00191720	zinc-finger,	gi|229595569	44	3	14.7
TTHERM_00723490	drug/metabolite transporter (DMT) superfamily	gi|118399001	74	5	14.8
TTHERM_00971790	-	gi|118385953	75	5	15
TTHERM_00047170	RRM, nucleolar phosphoprotein	gi|146165272	30	2	15
TTHERM_00537420	FOP N terminal dimerisation domain	gi|118380777	45	3	15
TTHERM_00382110	myb-like DNA-binding domain	gi|118364547	45	3	15
TTHERM_01043060	U1 small nuclear ribonucleoprotein A, splicing	gi|118396628	32	2	16
TTHERM_00614830	Dual specificity phosphatase	gi|118382993	132	8	16.5
TTHERM_00294670	-	gi|118359722	86	5	17.2
TTHERM_01299680	-	gi|118384313	70	4	17.5
TTHERM_00433680	Pre-mRNA-splicing factor cwc15	gi|118356269	35	2	17.5
TTHERM_00471820	micronuclear linker histone, HMG box family protein	gi|118401491	71	4	17.8
TTHERM_00322910	EMG1 nucleolar protein/ribosome biogenesis protein NEP1	gi|118359505	36	2	18
TTHERM_00938840	weak similarity to DNA polymerase III, alpha subunit	gi|118397453	37	2	18.5
TTHERM_00729210	Protein phosphatase 2C containing protein	gi|118379077	95	5	19
TTHERM_00459230	Dmc1p	gi|118382143	40	2	20
TTHERM_00219320	-	gi|118374833	40	2	20
TTHERM_00773720	-	gi|229594564	40	2	20
TTHERM_00218480	Bystin family protein, rRNA processing protein ENP1	gi|118374665	61	3	20.3
TTHERM_00277470	acetyl-CoA acyltransferases family protein	gi|146171814	41	2	20.5
TTHERM_00637490	NOL1/NOP2/sun family, ribosomal RNA methyltransferase	gi|146185731	88	4	22
TTHERM_00347930	DUF2361	gi|118379565	45	2	22.5
TTHERM_00216160	-	gi|118374597	114	5	22.8
TTHERM_00058600	-	gi|146161681	46	2	23
TTHERM_00013620	weak similarity to chromosome segregation protein SMC	gi|118350120	94	4	23.5
TTHERM_00052450	similar to human KIAA0556, function unknown	gi|118362952	241	10	24.1
TTHERM_00703510	-	gi|229594310	49	2	24.5
TTHERM_01001490	pre-mRNA splicing factor CWC21	gi|118388805	49	2	24.5
TTHERM_00077380	weak similarity to chromosome segregation protein SMC	gi|118364916	150	6	25
TTHERM_00460480	-	gi|118371089	77	3	25.7
TTHERM_00155330	PX domain (phosphoinositide binding module)	gi|118352781	52	2	26
TTHERM_00693070	-	gi|229594419	52	2	26
TTHERM_00051720	K homology (KH) RNA-binding domain	gi|118363004	80	3	26.7
TTHERM_01020880	Serrate (RNA-processing multi-protein complex) RNA effector molecule	gi|229594069	54	2	27
TTHERM_00575520	Protein kinase, similar to MAPK2	gi|118352144	109	4	27.3

Appendix 1

gray: related to rRNA synthesis, transcription, splicing
pink: RdRP-related

Identified Proteins	homology/domains/other names	Accession #	MW	# peptides	MW/pep
TTHERM_00046640	K homology (KH) RNA-binding domain	gi|146165238	86	3	28.7
TTHERM_00283980	-	gi|146172048	86	3	28.7
TTHERM_00421020	RRM, bruno-like	gi|118402075	87	3	29
TTHERM_00467530	weak similarity to DNA polymerase III and SMC	gi|118383790	235	8	29.4
TTHERM_00394670	-	gi|118357379	59	2	29.5
TTHERM_00666270	-	gi|118373606	60	2	30
TTHERM_00502200	-	gi|146181170	60	2	30
TTHERM_00530690	Zinc knuckle family protein	gi|118400905	61	2	30.5
TTHERM_00693080	cyclin	gi|118380821	61	2	30.5
TTHERM_00006110	weak similarity to chromosome segregation protein SMC	gi|118349622	124	4	31
TTHERM_00193310	von Willebrand factor type A domain containing protein	gi|118367749	126	4	31.5
TTHERM_00941500	-	gi|118387612	192	6	32
TTHERM_00923050	protein kinase, TPR Domain	gi|146180576	132	4	33
TTHERM_01055650	Zinc finger?	gi|118388872	135	4	33.8
TTHERM_00693190	-	gi|118380843	170	5	34
TTHERM_00317470	-	gi|118376492	68	2	34
TTHERM_00354800	tubby-like protein, cell signaling?	gi|118354231	68	2	34
TTHERM_00703930	-	gi|118399478	309	9	34.3
TTHERM_00497940	-	gi|229593569	69	2	34.5
TTHERM_00529530	-	gi|118400685	70	2	35
TTHERM_00035190	weak similarity to chromosome segregation protein SMC?	gi|118346293	214	6	35.7
TTHERM_00317260	-	gi|118376450	110	3	36.7
TTHERM_00268140	-	gi|146166154	113	3	37.7
TTHERM_00852750	-	gi|118386324	155	4	38.8
TTHERM_00170510	Protein kinase domain containing protein	gi|118350859	79	2	39.5
TTHERM_00466120	Tesmin/TSO1-like CXC domain, Lin54, transcription factor	gi|118383712	120	3	40
TTHERM_00420790	EF hand family protein, similar to Neuronal calcium sensor 2	gi|229594649	85	2	42.5
TTHERM_01394310	Rdn3p, Nucleotidyltransferase similar to Rdn1p	gi|118394510	85	2	42.5
TTHERM_00660330	-	gi|118373807	89	2	44.5
TTHERM_00734110	-	gi|118356607	89	2	44.5
TTHERM_00129040	ribosomal protein S6 kinase	gi|118366333	178	4	44.5
TTHERM_00326680	TPR Domain, RNA polymerase II associated protein 3,	gi|118386627	90	2	45
TTHERM_00312320	AAA domain (ATP-binding)	gi|118349073	142	3	47.3
TTHERM_01053030	CCT motif family protein	gi|118358510	142	3	47.3
TTHERM_00343950	-	gi|118370478	145	3	48.3
TTHERM_00849280	short similarity to Polyadenylation factor I complex, subunit FIP1	gi|118372066	102	2	51
TTHERM_00643510	glutamate receptor, Ligand-gated ion channel family protein	gi|118364296	103	2	51.5
TTHERM_00420670	TPR Domain containing protein	gi|146186287	104	2	52
TTHERM_00637030	weak similarity to chromosome segregation protein SMC	gi|146185711	110	2	55
TTHERM_01044630	-	gi|118396595	114	2	57
TTHERM_00492370	protein kinase, similar to MAP kinase	gi|118379929	352	6	58.7
TTHERM_00046690	weak similarity to chromosome segregation protein SMC	gi|118363310	124	2	62
TTHERM_00790940	Actin and Kinesin motor domain fused	gi|118386141	125	2	62.5
TTHERM_00131240	DEAD/DEAH box helicase family protein	gi|118372814	127	2	63.5
TTHERM_01091290	CnjBp, Zinc knuckle family protein	gi|146182859	200	3	66.7
TTHERM_00077160	-	gi|118364872	275	4	68.8
TTHERM_00637010	weak similarity to SMC, and Mg2+ and Co2+ transporters	gi|118400078	145	2	72.5
TTHERM_00257250	-	gi|118371964	146	2	73

Appendix 1

gray: related to rRNA synthesis, transcription, splicing
pink: RdRP-related

Identified Proteins	homology/domains/other names	Accession #	MW	# peptides	MW/pep
TTHERM_00131340	weak similarity to phosphate acyltransferase	gi|146175974	225	3	75
TTHERM_00691160	-	gi|118399513	155	2	77.5
TTHERM_01216140	Dcr1p, Dicer-related protein	gi|118352386	247	3	82.3
TTHERM_01094800	-	gi|118355304	176	2	88
TTHERM_00703320	-	gi|118399356	381	4	95.3
TTHERM_00420400	zinc-finger?	gi|146186273	194	2	97
TTHERM_00497790	Eukaryotic rRNA processing protein EBP2	gi|118389844	310	3	103.3
TTHERM_01232260	-	gi|118385876	366	3	122
TTHERM_00670640	PHD-finger family protein	gi|118386533	437	3	145.7
TTHERM_00647140	-	gi|118388654	647	4	161.8