A SUMO E3 ligase promotes long non-coding RNA transcription to regulate small RNA-directed DNA elimination

Salman Shehzada†, Tomoko Noto, Julie Saksouk & Kazufumi Mochizuki*

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

†Present address: Department of Genetic Medicine and Development, University of Geneva, Geneva, Switzerland

*Correspondence: kazufumi.mochizuki@cnrs.fr

Abstract

Small RNAs target their complementary chromatin regions for gene silencing through nascent long non-coding RNAs (lncRNAs). In the ciliated protozoan *Tetrahymena*, the interaction between Piwi-associated small RNAs (scnRNAs) and the nascent lncRNA transcripts from the somatic genome has been proposed to induce target-directed scnRNA degradation (TDSD), and scnRNAs not targeted for TDSD later target the germline-limited sequences for programmed DNA elimination. In this study, we show that the SUMO E3 ligase Ema2 is required for the accumulation of lncRNAs from the somatic genome and thus for TDSD and completing DNA elimination to make viable sexual progeny. Ema2 interacts with the SUMO E2 conjugating enzyme Ubc9 and enhances SUMOylation of the transcription regulator Spt6. We further show that Ema2 promotes the association of Spt6 and RNA polymerase II with chromatin. These results suggest that Ema2-directed SUMOylation actively promotes lncRNA transcription, which is a prerequisite for communication between the genome and small RNAs.

Introduction

Small RNAs of approximately 20-30 nucleotides that are complexed with Argonaute family proteins target either mRNAs for post-transcriptional silencing or chromatin regions for transcriptional gene silencing (Holoch & Moazed, 2015; Wilson & Doudna, 2013). For the latter process, small RNAs are generally considered to recognize their genomic targets via nascent long non-coding RNAs (lncRNAs) to induce heterochromatin formation. Therefore, chromatin regions that are targeted for silencing by small RNAs must be paradoxically transcribed to provide nascent lncRNAs

In fission yeast, small interfering RNAs (siRNAs) mediate the deposition of histone 3 lysine 9 (H3K9me) for heterochromatin assembly at centromeric repeats (Volpe *et al*, 2002; Hall *et al*, 2002). While the HP1 protein Swi6 binds to H3K9me for transcriptional silencing, phosphorylation of histone H3 serine 10 at the M phase of the cell cycle evicts Swi6, passively allowing lncRNA transcription from the centromeric repeats at G1 and S phase and thus promoting H3K9me deposition to newly assembled nucleosomes (Chen *et al*, 2008; Kloc *et al*, 2008). In contrast, the HP1 paralog Rhino in fruit flies, which is specifically enriched at PIWI-interacting RNA (piRNA) clusters by binding to H3K9me, actively recruits dedicated variants of basal transcription factors, allowing lncRNA transcription from heterochromatin while preventing mRNA transcription from transposons in the same loci (Andersen *et al*, 2017). A similar active heterochromatin-dependent lncRNA transcription has been reported in plants, where SHH1, a reader of H3K9me, recruits the plant-specific RNA polymerase IV (Law *et al*, 2013, 2011).

Because small RNA-producing loci are also small RNA targets, it is difficult to separately study lncRNA transcription for small RNA biogenesis and that for small RNA-dependent recruitment of downstream effectors. Therefore, it remains unclear whether the biogenesis and reception of small RNAs share the same lncRNA transcription mechanism in small RNA-directed heterochromatin formation processes. In contrast, the source and target loci of small RNAs reside in different nuclei during programmed DNA elimination in the unicellular eukaryote *Tetrahymena thermophila*, which provides a unique system to study the mechanism and roles of lncRNA transcription in small RNA-directed chromatin regulation.

Like most other ciliates, each *Tetrahymena* cell contains two types of nuclei, the diploid germline micronucleus (MIC) and the polyploid somatic macronucleus (MAC). During conjugation, the sexual reproduction process of *Tetrahymena*, the MIC undergoes meiosis and fertilization, followed by the formation of the new MIC and MAC, while the parental MAC

is degraded (Chalker *et al*, 2013). When the new MAC develops, the 200 Mb MIC genome is downsized to the 103 Mb MAC genome by programmed DNA elimination. This process removes ~12,000 internal elimination sequences (IESs), many of which include transposons, and the remaining MAC-destined sequences (MDSs) are re-ligated to form the MAC chromosomes (Noto & Mochizuki, 2017).

DNA elimination in *Tetrahymena* is regulated by three types of lncRNAs, which occur in different nuclei at distinct times (Aronica *et al*, 2008; Schoeberl & Mochizuki, 2011). During meiotic prophase, lncRNAs are transcribed bidirectionally in the MIC in a genome-wide manner by RNA polymerase II and dedicated conjugation-specific Mediator-associated proteins (Chalker & Yao, 2001; Schoeberl *et al*, 2012; Mochizuki & Gorovsky, 2004b; Tian *et al*, 2019; Garg *et al*, 2019). These MIC-lncRNAs are processed into ~29-nt small RNAs, called scnRNAs, by a Dicer homolog (Malone *et al*, 2005; Mochizuki & Gorovsky, 2005) and loaded into the Piwi-clade Argonaute protein Twi1 (Mochizuki *et al*, 2002; Noto *et al*, 2010).

The Twi1-scnRNA complex then moves into the parental MAC, where pMAC-ncRNAs are transcribed bidirectionally during the mid-conjugation stages (Woo *et al*, 2016). The RNA helicase Ema1 promotes the interaction between pMAC-lncRNAs and scnRNAs (Aronica *et al*, 2008) and this interaction induces target-directed scnRNA degradation (TDSD), resulting in selective retention of IES-derived scnRNAs (Aronica *et al*, 2008; Schoeberl *et al*, 2012; Mochizuki & Gorovsky, 2004a; Noto *et al*, 2015; Noto & Mochizuki, 2018). A similar TDSD has also been suggested in *Paramecium*, and the importance of pMAC-lncRNAs for DNA elimination was demonstrated by disrupting certain pMAC-lncRNAs by RNAi in this ciliate (Lepère *et al*, 2008). Trimethylation of histone 3 lysine 27 (H3K27me3) also occurs in the parental MAC in a scnRNA- and Polycomb repressive complex 2 (PRC2)-dependent manner (Liu *et al*, 2007), although the role of H3K27me3 in the parental MAC, if any, is unclear.

When the new MAC develops, the retained Twi1-scnRNA complexes translocate to the new MAC, where yet another type of lncRNA, nMAC-lncRNAs, is transcribed. The interaction between nMAC-lncRNA and scnRNA, which is also dependent on Ema1 (Aronica *et al*, 2008), is believed to recruit PRC2, which catalyzes both H3K9me3 and H3K27me3 for IES-specific heterochromatin assembly (Liu *et al*, 2007; Xu *et al*, 2021), and facilitate the secondary production of scnRNAs, which further promotes heterochromatin assembly (Noto *et al*, 2015). Heterochromatinized IESs are eventually excised by a domesticated transposase (Cheng *et al*, 2010; Vogt & Mochizuki, 2013).

Among the above three lncRNAs, pMAC-lncRNA does not produce small RNAs and is believed to be specialized for receptor function. Therefore, pMAC-lncRNA transcription and the following TDSD in the parental MAC in *Tetrahymena* provides a unique paradigm to investigate how the genome is transcribed to communicate with small RNAs. In this study, we show that Ema2, the conjugation-specific E3 ligase for small ubiquitin-like modifier (SUMO), is required for pMAC-lncRNA transcription and thus provides a tool to dissect the role of and the molecular mechanism for the transcription of this lncRNA.

Results

Ema2 is required for completing DNA elimination

EMA2 (TTHERM_00113330) mRNA is exclusively expressed during conjugation (Figure 1A). The encoded Ema2 protein tagged with HA at the endogenous *EMA2* locus was not detectable in the vegetative cells (Figure 1B, Vg) and first appeared in the MAC during conjugation (Figure 1B, 3 hours post-induction of mating [hpm] and 6 hpm). At the onset of new MAC development, Ema2 disappeared from the parental MAC and appeared in the new MAC (Figure 1B, 8 hpm), which later faded away (Figure 1B, 12 and 14 hpm). The conjugation-specific expression and the localization switch from the parental to the new MAC are reminiscent of the factors involved in DNA elimination (Mochizuki *et al*, 2002; Coyne *et al*, 1999; Kataoka & Mochizuki, 2015; Liu *et al*, 2007; Yao *et al*, 2007).

DNA elimination of exconjugants (sexual progeny) at 36 hpm was analyzed by DNA fluorescent in situ hybridization (FISH) using probes complementary to the transposable element Tlr1 (Wuitschick *et al*, 2002). DNA elimination is completed by ~14-18 hpm in wild-type cells (Austerberry *et al*, 1984; Mutazono *et al*, 2019), and the Tlr1 element was detected only in the MICs in the exconjugants from wild-type cells (Figure 2A, WT). In contrast, the Tlr1 element was detected in both the MICs and the MACs in the exconjugants from the *EMA2* somatic KO strains, in which all *EMA2* copies in the MAC were disrupted (Shehzada & Mochizuki, 2022) (Figure 2A, KO). The intensity of the FISH signal in the new MACs was lower in the exconjugants from the *EMA2* KO cells than in those from the *TWI1* KO cells (Figure 2B), in the latter of which DNA elimination is known to be completely blocked (Noto *et al*, 2015). Therefore, DNA elimination was partially blocked in the exconjugants of *EMA2* KO cells. Consistent with the requirement of DNA elimination in the viability of sexual progeny (Cheng *et al*, 2010; Vogt & Mochizuki, 2013), *EMA2* KO cells did not produce viable progeny (Figure 2C). Altogether, we conclude that maternally expressed *EMA2* is required for completing DNA elimination.

Ema2 is required for target-directed scnRNA degradation (TDSD)

We next asked whether *EMA2* is involved in TDSD. The production of scnRNAs occurs from IESs and their surrounding MDS regions in the MIC at the early conjugation stages (~2-3 hpm), and scnRNAs complementary to the MAC genome (= MDSs) are subjected to TDSD in the parental MAC in the mid-conjugation stages (~3-7 hpm), resulting in the selective retention of IES-derived scnRNAs that later target IESs for DNA elimination in the new MAC

(Mochizuki & Gorovsky, 2004a; Aronica *et al*, 2008; Schoeberl *et al*, 2012; Noto *et al*, 2015). We first compared scnRNAs at different conjugation stages by northern blot analysis using the 50-nt Mi-9 probe that is complementary to a repetitive sequence found in both IESs and MDSs (Aronica *et al*, 2008). Because scnRNAs detected by this probe are complementary to the MAC genome, in wild-type cells, the Mi-9-complementary scnRNAs were detected at 3 hpm, reduced at 4.5 hpm, and became undetectable at 6 hpm due to TDSD (Figure 3A, WT). In contrast, the Mi-9-complementary scnRNAs remained detectable at 6 hpm in the *EMA2* KO cells (Figure 3A, KO), suggesting that Ema2 is required for TDSD.

We also analyzed TDSD by small RNA sequencing at 3 and 8 hpm. Sequenced 26- to 32-nt small RNAs, which correspond to scnRNAs (Schoeberl *et al*, 2012), were mapped to the genomic tiles of MDSs and IESs (see Materials and Methods). Although scnRNAs that mapped many of the MDS tiles were greatly reduced from 3 to 8 hpm due to TDSD in the wild-type cells (Figure 3B, left, WT), those in the *EMA2* KO cells were only slightly reduced by 8 hpm (Figure 3B, left, *EMA2* KO). This TDSD defect in the *EMA2* KO cells seems milder than that in the *EMA1* KO cells, in which the amount of MDS-complementary scnRNAs remained constant until 8 hpm (Figure 3B, left, *EMA1* KO), indicating that Ema1 and Ema2 act differently in TDSD. In contrast, the amounts of scnRNAs complementary to IES tiles at 3 hpm and 8 hpm were comparable in the strains (Figure 3B, right), which is consistent with previous observations that only scnRNAs complementary to MDSs are targeted for TDSD. These results suggest that Ema2 is required for TDSD at the genome-wide level. Because it is known that loss of TDSD results in a partial block of DNA elimination (Aronica *et al*, 2008), the mild DNA elimination defect in *EMA2* KO cells (Figure 2B) can be explained by the requirement of Ema2 in TDSD.

Ema2 is required for the accumulation of SUMOylated proteins during conjugation

Ema2 possesses an SP-RING domain that has been found in many SUMO E3 ligases (Hochstrasser, 2001). The SP-RING domain of Ema2 is atypical (Figure 4A) in that the first cysteine of the zinc ion-binding residues in typical counterparts is replaced by histidine and some of the stabilizer residues (Duan *et al*, 2009; Yunus & Lima, 2009) are likely missing. Nonetheless, we found that Ema2 can directly interact with the *Tetrahymena* SUMO E2 enzyme Ubc9 in vitro (Figure 4B), suggesting that Ema2 is a bona fide SUMO E3 ligase.

To examine the role of Ema2 in SUMOylation, we expressed HA-tagged Smt3 (HA-Smt3) in an *EMA2* KO strain. SUMO is solely encoded by *SMT3* in *Tetrahymena* (Nasir *et al*, 2015), and HA-Smt3 could replace the essential function of endogenous Smt3 (Supplementary

Figure 1). We then crossed this strain with either a wild-type strain (called the WT cross) or another *EMA2* KO strain (called the *EMA2* KO cross). Because proteins and mRNAs are exchanged between two mating pairs through the conjugation junction, *EMA2* mRNA/Ema2 protein expressed in the wild-type partner of the WT cross is expected to move into the *EMA2* KO partner and restore the *EMA2* KO phenotypes.

Total proteins were harvested at 4.5 and 6 hpm, and SUMOylated proteins were detected by western blotting using an anti-HA antibody (Figure 4C). At both time points, high molecular weight proteins (mainly >200 kDa) were detected in the WT cross (Figure 4C, WT), and they were reduced to ~50% in the *EMA2* KO cross (Figure 4C, KO). These results indicate that Ema2 is the major SUMO E3 ligase during the mid-conjugation stages. The requirement of protein SUMOylation in DNA elimination was previously demonstrated in *Paramecium* by RNAi knockdown of *UBA2* and *SUMO* (Matsuda & Forney, 2006). Therefore, involvement of a SUMO pathway in DNA elimination is likely conserved among ciliates.

Ema2 is required for SUMOylation of Spt6

Next, to identify the SUMOylation target(s) of Ema2, we introduced a construct expressing His-tagged Smt3, which can also replace the essential function of Smt3 (Supplementary Figure 1), into an *EMA2* KO strain and crossed it with a wild-type strain (WT cross) or another *EMA2* KO strain (KO cross). Then, SUMOylated proteins at 6 hpm were concentrated using a nickel column and identified by mass spectrometry. Although most of the proteins were detected similarly between the WT cross and *EMA2* KO cross, Spt6 was detected more abundantly from the WT cross (Figure 5A).

To confirm the Ema2-dependent SUMOylation of Spt6, we introduced a construct expressing HA-tagged Spt6 (Spt6-HA) from the endogenous *SPT6* locus into an *EMA2* KO strain and mated it with a wild-type strain (WT cross) or another *EMA2* KO strain (*EMA2* KO cross). Then, total proteins harvested at 4.5 and 6 hpm were analyzed by western blotting using an anti-HA antibody (Figure 5B). In the WT cross, a slower migrating population of Spt6-HA was detected in addition to a band corresponding to unmodified Spt6-HA at both time points (Figure 5B, WT). In contrast, slower migrating Spt6-HA was barely detectable in the *EMA2* KO cross (Figure 5B, KO).

The nature of slower migrating species of Spt6-HA was further examined by immunoprecipitating Spt6-HA using an anti-HA antibody. Among the total purified Spt6-HA (Figure 5C, left, IP-WT), the slower migrating species were detected by an anti-Smt3 antibody in the WT cross (Figure 5C, right, IP-WT), and such SUMOylated Spt6-HA species were

greatly reduced in the *EMA2* KO cross (Figure 5C, KO). These results indicate that the slower migrating species of Spt6 are SUMOylated Spt6 and Ema2 is required for the majority of SUMOylation of Spt6 during the mid-conjugation stages.

Ema2 is required for the accumulation of lncRNA in the parental MAC

Spt6 is a conserved regulator of several steps of transcription in various eukaryotes. Because TDSD was proposed to be triggered by the base-pairing interaction between scnRNAs and nascent pMAC-lncRNAs in the parental MAC (Aronica *et al*, 2008; Noto & Mochizuki, 2018), we hypothesized that Ema2-dependent Spt6 SUMOylation promotes pMAC-lncRNA transcription.

To examine pMAC-lncRNAs, we amplified transcripts spanning IES-MDS borders by RT–PCR in which MAC-lncRNAs can be distinguished from MIC-lncRNAs by their lengths (Figure 6A). For the three loci examined, pMAC-lncRNAs were detected in wild-type cells but not in *EMA2* KO cells at 6 hpm, although the control *RPL21* mRNA was detected in both conditions (Figure 6A). The lack of detection of pMAC-lncRNAs in *EMA2* KO cells was not due to loss of the primer binding sites by alternative DNA elimination in their prior sexual reproductions, as genomic PCR with the same primer sets detected the corresponding MAC loci in the *EMA2* KO cells (Supplementary Figure 2). We therefore conclude that Ema2 is required for the accumulation of pMAC-lncRNAs at least for the three tested loci.

The lncRNA accumulation in the parental MAC was also examined cytologically using the J2 antibody, which specifically detects long double-stranded (ds)RNAs (Schönborn *et al*, 1991). The lncRNA transcription in the three different nuclei occurs bidirectionally (Woo *et al*, 2016; Mochizuki & Gorovsky, 2005; Malone *et al*, 2005; Aronica *et al*, 2008), and thus, in wild-type cells (Figure 6B, WT), long dsRNAs were first detected in the MIC at its premeiotic stage (3 hpm), then found in the parental MAC at the mid stages of conjugation (6 hpm), and finally detected in the new MAC at late conjugation (9 hpm). In the *EMA2* KO cells, long dsRNAs were detected in the MIC at 3 hpm and in the new MAC at 9 hpm but were undetectable in the parental MAC at 6 hpm (Figure 6B, *EMA2* KO), indicating that Ema2 is globally required for the accumulation of pMAC-lncRNAs.

In contrast, *EMA1* was dispensable for the accumulation of dsRNAs in the parental MAC (Figure 6B, *EMA*1 KO), which is consistent with the previous notions that Ema1 is required for pMAC-lncRNAs to interact with scnRNAs but not their transcription (Aronica *et al*, 2008). The milder TDSD defect in the *EMA2* KO cells than in the *EMA1* KO cells (Figure

3B) indicates that some Ema1-dependent TDSD still occurs in the absence of Ema2, possibly through Ema2-independent transcription of mRNAs and/or ncRNAs.

In parallel to TDSD, scnRNA-dependent accumulation of H3K27me3, which is catalyzed by the histone methyltransferase Ezl1 within PRC2, occurs in the parental MAC (Liu *et al*, 2007). We found that H3K27me3 was undetectable in the parental MAC in the *EMA2* KO cells at 6 hpm or in the *EMA1*, *TWI1* and *EZL1* KO cells (Figure 6C). Therefore, Ema2 is required for the scnRNA-directed deposition of H3K27me3 in the parental MAC, which can be explained by the loss of the chromatin recruitment of PRC2 by the scnRNA-lncRNA interaction. Altogether, these results indicate that Ema2 is required for the genome-wide accumulation of pMAC-lncRNAs.

Ema2 facilitates chromatin association of Spt6 and RNA polymerase II

Next, we compared the localization of Spt6 and Rpb3, the third largest subunit of RNA polymerase II (RNAPII). In both WT and *EMA2* KO cross described above, Spt6-HA and Rpb3 were similarly detected in the MAC at 4.5 hpm by immunofluorescence staining (Figure 7A). Therefore, at least at the cytological level, the absence of Ema2 does not affect the localization of Spt6 and RNAPII.

We then examined their subnuclear localization by cell fractionation (Figure 7B, left), in which cells at 4.5 hpm were first incubated with lysis buffer to extract soluble cyto- and nucleoplasmic proteins, and then, the insoluble fraction was sonicated to extract chromatin-bound proteins (Ali *et al*, 2018). In the WT cross (Figure 7B, WT), Spt6-HA was detected in both the soluble (S1) and chromatin-bound (S2) fractions, but slower-migrating SUMOylated Spt6-HA was detected only in the latter fraction, indicating that SUMOylated Spt6 is associated with chromatin. A similar distribution was detected for Rpb3. This chromatin association of Spt6 and RNAPII does not require RNA, as they were retained in the insoluble fraction after RNase A treatment without sonication (Figure 7C), although Twi1, which interacts with chromatin through nascent pMAC-lncRNA (Arnica et al 2008), was mostly detected in the soluble fraction after the treatment (Figure 7C). In the *EMA2* KO cross, Spt6-HA in the chromatin-bound fraction was greatly reduced (Figure 7B, KO, S2). Furthermore, much less chromatin-bound Rpb3 was detected in the *EMA2* KO cross than in the WT cross (Figure 7B, Rpb3). We therefore conclude that Ema2 promotes the association of Spt6 and RNAPII with chromatin in the parental MAC at the mid-conjugation stage.

Ema2 promotes lncRNA transcription independently of Spt6 SUMOylation

Next, to directly examine the importance of SUMOylation of Spt6, we intended to produce an unSUMOylateable Spt6 mutant. *Tetrahymena* Spt6 consists of an N-terminal domain of low complexity (DOL) and a downstream region that is conserved among eukaryotes (Figure 8A, WT). We produced constructs expressing HA-tagged Spt6 with lysine to arginine substitutions either for all lysine residues in the DOL (*HA-SPT6-DOL-KR*) or for lysine residues that were on the surface of an AlphaFold predicted *Tetrahymena* Spt6 structure in one of the three non-DOL parts (N, M and C in Figure 8A) in addition to those that were also predicted to be SUMOylation targets by the JASSA algorithm in all three parts (Beauclair *et al*, 2015) (*HA-SPT6-N-KR, HA-SPT6-M-KR and HA-SPT6-C-KR*, Figure 8A).

We introduced these mutant constructs as well as the construct expressing wild-type Spt6 (*HA-SPT6-WT)* into *SPT6* KO cells and found that all of them were able to restore the lethality of *SPT6* KO cells (Supplementary Figure 3). However, the cells rescued by *HA-SPT6-N-KR* and *HA-SPT6-M-KR* showed severe defects in meiotic progression and mating initiation, respectively, making their SUMOylation status during conjugation uninvestigable. In contrast, the cells rescued by *HA-SPT6-DOL-KR* and *HA-SPT6-C-KR* showed a seemingly normal progression of conjugation, although the *HA-SPT6-C-KR* cells showed lower mating efficiency. We found that while HA-Spt6-DOL-KR was SUMOylated like HA-Spt6-WT, SUMOylation was not detected on HA-Spt6-C-KR (Figure 8B). Therefore, Spt6-C-KR is an unSUMOylatable Spt6 mutant.

Cell fractionation experiments showed that HA-Spt6 and RNAPII were detected in the chromatin-bound fraction in *HA-SPT6-C-KR* cells as in the *HA-SPT6-WT* cells (Figure 8C), and long dsRNAs were detected in the parental MAC in the *HA-SPT6-C-KR* cells (Figure 8D). These results suggest that, contrary to our expectation, Spt6 SUMOylation per se is not required for the lncRNA transcription in the parental MAC. Nonetheless, exconjugants from the *HA-SPT6-C-KR* cells showed a mild DNA elimination defect (Figure 8E), indicating that Ema2-directed Spt6-SUMOylation plays a role in DNA elimination other than promoting pMAC-lncRNA transcription.

Discussion

In this study, we showed that the conjugation-specific SUMO E3 ligase Ema2 is required for the accumulation of lncRNAs (Figure 6), TDSD (Figure 3A, B) and heterochromatin formation (Figure 3C) in the parental MAC and eventually for completing DNA elimination (Figure 2) in *Tetrahymena.* We found that Ema2 is responsible for SUMOylation of the transcriptional regulator Spt6 (Figure 5) and promotes the interaction of Spt6 and RNAPII with chromatin (Figure 7). We therefore conclude that Ema2 facilitates genome-wide lncRNA transcription in the parental MAC, which is a prerequisite for scnRNA-chromatin communication and thus for the downstream TDSD that regulates DNA elimination.

Our observation that Ema2 enhances the chromatin interaction of Spt6 and RNAPII (Figure 7) seems to contradict the fact that Spt6 and RNAPII are essential for vegetative cell viability (Supplementary Figure 3) (Mochizuki & Gorovsky, 2004b), but Ema2 is expressed exclusively during conjugation (Figure 1). Therefore, the loss of the majority of Spt6 and RNAPII from chromatin in the absence of Ema2 (Figure 7B) must be a temporal event during conjugation. It is likely that some radical change in the chromatin environment in the parental MAC is required for the global transcription of pMAC-lncRNAs, and Ema2-dependent SUMOylation might maintain Spt6 and RNAPII on chromatin for transcription in such an environment. Alternatively, because the expression of many of the transcriptional machineries, including Spt6 and RNAPII subunits, is highly upregulated during conjugation (Miao *et al*, 2009; Xiong *et al*, 2012; Mochizuki & Gorovsky, 2004b), the transcription of pMAC-lncRNAs might require a greater amount of transcriptional machinery than other types of transcription, and Ema2-dependent SUMOylation might be required for committing these excess machineries to chromatin for transcription.

Our investigation of the unSUMOylatable Spt6 mutant suggested that although Spt6 is SUMOylated in an Ema2-dependent manner (Figure 5), SUMOylation of Spt6 by itself is not required for pMAC-lncRNA transcription (Figure 8). Therefore, Ema2 facilitates pMAC-lncRNA transcription likely through SUMOylation of some protein(s) other than Spt6. Although we failed to detect any other SUMOylated protein in an Ema2-dependent manner (Figure 5A), some Ema2-dependent SUMOylation events might not be detected in our current approach if Ema2-dependent and Ema2-independent SUMOylation occur in the same proteins and/or if Ema2-dependent SUMOylation results in protein degradation, such as through SUMO-targeted ubiquitination (Praefcke *et al*, 2012; Staudinger, 2017). Alternatively, SUMOylation of Spt6 may promote pMAC-lncRNA transcription by competing with some

other modification that occurs at the same lysine residue(s) and our unSUMOylatable Spt6 mutant might actually mimics SUMOylated state of Spt6 by preventing such modification. To further investigate the role of Ema2, future studies will need to identify individual Ema2-dependent SUMOylated lysine residues under conditions of in vivo proteasome inhibition and to compare the status of other modifications of Spt6 in the presence and absence of Ema2. These are also important to understand the role of Spt6 SUMOylation in DNA elimination (Figure 8E).

It was reported that SUMOylation of the heterochromatin components Swi6, Chp2 and Clr4 is important for RNAi-directed heterochromatin silencing in fission yeast (Shin *et al*, 2005). Additionally, SUMOylation of the histone deacetylase HDAC1 promotes small RNA-directed transcriptional silencing in *C. elegans* (Kim *et al*, 2021). Furthermore, in the piRNA-directed transposon silencing of *Drosophila*, the SUMO E3 ligase Su(var)2-10 plays an essential role in the recruitment of the histone methyltransferase complex SetDB1, which deposits H3K9me3 (Ninova *et al*, 2020a, 2020b), and Su(var)2-10-independent SUMOylation of Panoramix (Panx) also promotes transcriptional silencing by recruiting the general heterochromatin effector Sov (Andreev *et al*, 2022). All these SUMOylation events seem to require pre-existing lncRNA transcription. In contrast, this study showed that Ema2-dependent SUMOylation during the process of DNA elimination in *Tetrahymena* occurs upstream of lncRNA transcription (Figure 6A, B). In addition, the fact that the loss of the Piwi protein Twi1 blocks the accumulation of H3K27me3 in the parental MAC (Liu *et al*, 2007; Xu *et al*, 2021) without affecting lncRNA accumulation (Woo *et al*, 2016) indicates that heterochromatin formation is not a prerequisite for lncRNA transcription in *Tetrahymena*, in contrast to heterochromatin-dependent lncRNA transcription of piRNA clusters in *Drosophila* (Andersen *et al*, 2017). Altogether, *Tetrahymena* likely uses a unique evolutionary solution for transcribing lncRNAs to promote target-directed degradation of small RNAs and small RNA-directed heterochromatin formation. We therefore believe that further investigations of the role of Ema2-dependent SUMOylation in lncRNA transcription will clarify an undescribed layer of regulatory mechanisms in small RNA-directed chromatin regulation.

Materials and Methods

Strains and cell culture condition

The *Tetrahymena thermophila* wild-type strains B2086, CU427 and CU428 and the MIC defective "star" strains B*VI and B*VII were obtained from the *Tetrahymena* stock center. The production of the *EMA2* KO strains were described previously (Shehzada & Mochizuki, 2022). The other transgenic *Tetrahymena* strains are described below. Cells were grown in SPP medium (GOROVSKY et al. 1975) containing 2% proteose peptone at 30 °C. For conjugation, growing cells (~5-7 x 105/mL) of two different mating types were washed, pre-starved (~12-24 hr) and mixed in 10 mM Tris (pH 7.5) at 30 °C.

Oligo DNA, Synthetic DNA and antibodies

The primers and synthetic DNAs used in this study are listed in Supplementary Information. The following antibodies were obtained from manufacturers: anti-H3K27me3 (Millipore, 07-449); anti-HA HA11 (Covance, clone 6B12); anti-H3K9me3 (Active Motif, 39162); anti-GST (BD Biosciences, 554805); anti-His (Proteintech, 66005-1-Ig); anti-long dsRNA J2 (Jena Bioscience RNT-SCI-10010200); anti-α-tubulin 12G10 (Developmental Studies Hybridoma Bank of the University of Iowa). The anti-Rpb3 antibody was described previously (Kataoka & Mochizuki, 2017). The anti-Smt3 antibody was a gift from Dr. James Forney (Purdue University) and described previously (Nasir *et al*, 2015).

Establishment of *EMA2*-HA strains

The C-terminal part of the *EMA2* coding sequence was amplified from the genomic DNA of B2086 using EMA2-C-FW and EMA2-HA-C-RV. Similarly, the 3' flanking sequence of *EMA2* was amplified using EMA2-F-FW and EMA2-F-RV. The two PCR products were combined by overlapping PCR using EMA2-C-FW and EMA2-F-RV and cloned into XbaI and XhoI sites of pBlueScriptSK (+) to obtain pEMA2. Next, the *HA-neo3* construct was excised out from pHA-neo3 (Kataoka *et al*, 2010) using SalI and BamHI and cloned into the same restriction sites of pEMA2 to obtain pEMA2-HA-neo3. Finally, pEMA2-HA-neo3 were digested with XbaI and XhoI and introduced into the *EMA2* MAC locus of B2086 and CU428 by homologous recombination using biolistic transformation (Bruns & Cassidy-Hanley, 2000) followed by the selection with 100 μg/mL paromomycin and 1 μg/mL $CdCl_2$. Then, the endogenous *EMA2* MAC copies were replaced by the *EMA2-HA-neo3* locus through phenotypic assortment (Hamilton & Orias, 2000).

Immunofluorescent staining and western blotting

Immunofluorescent staining of long dsRNA was performed as described (Woo *et al*, 2016). The other immunofluorescent staining experiments were performed as described (Loidl & Scherthan, 2004). For immunofluorescent staining, all the primary and respective secondary antibodies (Alexa-488-coupled anti-rabbit or mouse IgG, Invitrogen) were diluted 1:1,000. For western blotting, all the primary and respective secondary antibodies (HRP-coupled anti-rabbit or mouse IgG, Jackson ImmunoResearch Lab) were diluted 1:10,000.

DNA elimination assay by FISH

DNA elimination assay by FISH was performed as described (Loidl & Scherthan, 2004). The plasmid DNAs pMBR 4C1, pMBR 2 and Tlr IntB, which contain different parts of the Tlr1-element were used to produce Cy5-labeled FISH probes by nick translation (Wuitschick *et al*, 2002). Images of stained cells were analyzed by Fiji software. Whole areas of the new MAC (A) and MIC (I) and a part of the cytoplasm (C) were manually encircled in the images; and the intensities of DAPI and Cy5 (FISH) signals in each area were measured. Then, the IES retention index was calculated as follows: 0.35 × {[(mean Cy3 intensity A - mean Cy3 intensity C)]/[(mean Cy3 intensity I - mean Cy3 intensity C)]}/{[(mean DAPI intensity A - mean DAPI intensity C)]/[(mean DAPI intensity I - mean DAPI intensity C)]}. Based on our previous observation, DNA elimination is completely blocked in the absence of *TWI1* KO therefore, the correction factor of 0.35 was obtained experimentally to make the average IES retention index of *TWI1* KO cells to 1.

Viability test for progeny

To test the progeny viability, mating pairs were isolated into drops of 1x SPP at 6 hpm and incubate for ~3 days. Then, completion of the conjugation of grown up cells from wild-type and *EMA2* KO pairs were examined for their 6-mp resistance and paromomycin sensitivity, respectively as described previously (Mochizuki *et al*, 2002).

Small RNA analyses

For northern hybridization analysis of scnRNAs, 10 μg of total RNAs isolated from conjugating wild-type and *EMA2* KO cells using TRIzol reagent were separated in a 15% acrylamide-urea gel and analyzed by northern blot using the radio-labeled Mi-9 probe as previously described (Aronica *et al*, 2008). The hybridized probe was detected by Typhoon

IP Phosphorimager (GE healthcare). High-throughput sequencing and analyses of small RNAs were performed as described (Mutazono et al 2019). The 2020 version of the MAC genome assembly (Sheng et al. 2020) was fragmented into 10 kb pieces. Each 10 kb fragment that contains longer than 3 kb mappable sequence was used as a MDS tile (total 10,235 tiles). Each Type-A IES (Noto *et al*, 2015) was used as an IES tile (total of 4,691 tiles). Normalized numbers (RPKM [reads per kilobase of unique sequences per million]) of 26- to 32-nt small RNAs that uniquely matched one of the MDS and IES tiles were counted.

GST pull down assay

E. coli codon-optimized synthetic *EMA2* and *UBC9* genes (*EMA2-Ec and UBC9-Ec*) were cloned into the BamHI and XhoI sites of pGEX-4T-1 and NdeI and BamHI sites of pET28a (+), respectively to obtain pGEX-4T-1-EMA2-Ec and pET28-UBC9-Ec. pET28-UBC9-Ec was introduced into BL21 *E. coli* strain, His-Ubc9 expression was induced ~16 hours at 25 °C with 0.05 mM isopropylthio-β-galactoside (IPTG), the cells were lysed in HEPES buffer (50 mM HEPES pH 7.5, 800 mM KCL, 10% glycerol, 0.2 mM PMSF, 1x cOmplete ETDA-free [Roche]) including 10 mM imidazole, incubated with Ni-NTA beads for 2 hours at 4 °C, washed four times with HEPES buffer including 20 mM imidazole, eluted with HEPES buffer including 250 mM imidazole, dialyzed with 50 mM Tris pH 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT, 50% glycerol and finally passed through Microcon 10 kDa filter (Merck). For GST pull-down assay, BL21 *E. coli* cells were transformed with pGEX-4T-1 or pGEX-4T-1-*EMA2-Ec*, and expression of GST-alone or GST-Ema2 was induced ~16 hours at 25 °C, lysed with 50 mM Tris–HCl, pH 7.5, 100 mM NaCl, 1 mM DTT, 0.2 mM PMSF,1x cOmplete EDTA-free and incubated with Glutathione Sepharose 4B for 90 min at 4 °C. Then, the beads were incubated with 10 µg of His-Ubc9 in GST-pull down buffer (20 mM Tris–HCl, pH 7.5, 100 mM NaCl, 1% NP-40, 5 mM $MgCl_2$, 1 mM DTT, 0.2 mM PMSF, 1x cOmplete EDTA-free) for 90 min at 4°C, washed 4 times with GST-pull down buffer and eluted with 2x SDS sample buffer. The eluted proteins were analyzed by western blotting.

Functionality test of HA-SMT3 and His-SMT3 using *SMT3* germline KO strains

The *SMT3* homozygous heterokaryon KO strain GC6 (Nasir *et al*, 2015) was obtained from Dr. James Forney (Purdue University). This strain was crossed with B2086, and sexual progeny were selected for their paromomycin resistance. After sexual maturation, clones showing parmomycin sensitivity (i.e. the *SMT3* KO locus was removed from the MAC by phenotypic assortment) were established as *SMT3* heterozygous heterokaryon KO strains. Then they were

crossed with B*VII to obtain their round I exconjugants. Finally, their genotypes were examined by PCR to select two *SMT3* homozygous heterokaryon KO strains, B1-1 and B5-17. They were then mated and pBCMB1-His-SMT3 (or pBCMB1-His for negative control) or pBCS3B1-HA-SMT3 (or pBCS3B1-HA for negative control) were introduced into the new MAC at 8 hpm by particle gun. The cells were incubated in 10 mM Tris pH 7.5 at 30 ºC until 24 hpm and fed by adding the equal volume of 2xSPP. After 3 hr incubation at 30 ºC, the *neo* expression (and His-Smt3 expression for pBCMB1-His-SMT3) was induced by adding 1 µg/mL CdCl2 and cells were incubated additional 1 hr. Then 100 µg/mL paromomycin was added and aliquoted to 96 well plates (150 µL/well, 7 plates). The plates were incubated at 30 ºC for 4 to 5 days and paromomysin resistant cells were further examined by genomic PCR to confirm genetic rescue by the introduced contracts (but not by the endogenous *SMT3* copy from the parental MAC).

Establishment of HA-Smt3 and His-Smt3 expressing strains

STM3 cDNA was amplified by RT-PCR from the total RNA of vegetatively growing CU428 using SMT3_HA_FW and SMT3_HisExt_RV and cloned into the BamHI and SpeI sites of pBNMB1-HA using NEBuilder HiFi DNA assembly kit to obtain pBNMB1-HA-SMT3. Next, neo5 was excised with SalI and XmaI and replaced with pur6, which was amplified from pBP6MB1 with BTU1-LCFW and MTT1_5UTR-SeqRV, using NEBuilder HiFi DNA assembly kit to obtain pBP6MB1-HA-SMT3. For His-SMT3 expression, the HA-tag encoding sequence of pBP6MB1-HA-SMT3 was replaced by a 6xHis tag-encoding sequence. Then, pBP6MB1-HA-SMT3 and pBP6MB1-His-SMT3 were digested by XhoI and introduced into the MAC BTU1 locus of the EMA2 KO cells. Transformation, selection, and phenotypic assortment were performed as described for the SPT6-HA strains.

Affinity purification of SUMOylated proteins

For purification of His-Smt3-conjugated proteins, *EMA2* KO cells containing the BP6MB1-His-SMT3 construct was mated with either a wild-type or another *EMA2* KO strain and His-SMT3 was induced by adding 0.05 µg/mL $CdCl_2$. The purification was performed as descried (Liebelt *et al*, 2019; Hendriks & Vertegaal, 2016). Briefly, the cells were lysed in Guanidine Lysis Buffer (6 M Guanidine-HCl, 93.2 mM Na_2HPO_4, 6.8 mM NaH_2PO_4, 10 mM Tris-HCl pH 8.0) by sonication (2 x 25 times, 20% power, 60% pulse) using a probe sonicator (Omni-Ruptor 250 Ultrasonic Homogenizer). After centrifugation, the cleared lysate was supplemented with 50 mM imidazole and 5 mM β-mercaptoethanol and incubated with Ni-

NTA agarose beads. The beads were washed, and the bound proteins were eluted with Elution buffer (6 M Urea, 58 mM Na_2HPO_4, 42 mM NaH_2PO_4, 10 mM Tris-HCl pH 8.0, 500 mM imidazole). The eluted proteins were precipitated in 10% TCA, resuspended in 1x SDS buffer, and analyzed by western blot and mass-spectrometry. For purification of HA-Smt3-conjugated proteins, total proteins were precipitated by incubating cells in 6 % trichloroacetic (TCA) on ice for 10 min and then spined down at 13,000 rpm for 5 min at 4 °C. The precipitate was dissolved in 1x SDS sample buffer, neutralized with 2 M Tris, and then incubated for 5 min at 95°C. The dissolved proteins (600 µl) were then suspended in 7.4 mL of 20 mM Tris pH 7.5, 100 mM NaCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$ and incubated with Ezview Red anti-HA agarose beads (Sigma) for overnight at 4 °C. Then the beads were washed 5 times with 20 mM Tris pH 7.5, 100 mM NaCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 0.1% tween 20 and the bound proteins were eluted by incubating with 1x SDS buffer for 5 min at 95°C. The eluted proteins were analyzed by western blotting.

Establishment of Spt6-HA expressing strains

To obtain the *SPT6*-HA construct, a part of the coding sequence of *SPT6* was amplified from the genomic DNA of B2086 using *SPT6*-FW1-SacII and *SPT6*-RV2-overlap. Similarly, the 3' flanking sequence was amplified using *SPT6*-3F-FW2-overlap and *SPT6*-3F-RV1-KpnI. The two PCR products were combined by overlapping PCR using *SPT6*-FW1-SacII and *SPT6*-3F-RV1-KpnI and cloned into the SacII and KpnI sites of pBlueScriptSK (+) to obtain p*SPT6*. Next, the NheI and XhoI fragment of pHA-neo3 was cloned into the same restriction sites of p*SPT6* to obtain p*SPT6-HA-neo3*. Then, *neo3* was excised from p*SPT6*-HA-neo3 with XmaI and SalI and replaced by *pur6* from pBP6MB1-TR-TUBE to obtain p*SPT6-HA-pur6*. Finally, the p*SPT6-HA-pur6* plasmid were digested by SacII and KpnI and introduced into the MAC *SPT6* locus of one of the *EMA2* KO strains by biolistic transformation. The transformed cells were selected in 300 µg/mL puromycin (Invitrogen) and 1 µg/mL $CdCl_2$ and cells were phenotypically assorted as previously described (Hamilton & Orias, 2000), until cells grew in 1.2 mg/mL of puromycin.

RT-PCR analyses of lncRNA transcripts

Total RNA from mating wild-type and *EMA2* KO cells were isolated at 6 hpm using TRIzol reagent (Invitrogen). The RNA was treated with Turbo DNase (Ambion) and cDNA was synthesized from 1 µg of RNA using SuperScript II (Invitrogen) with random 6-mer as primers. The lncRNA transcripts from the MAC loci containing the M-, L8-, and R2 IES

boundaries (see also Figure 6A) were amplified by nested PCR (95°C for 20sec/50°C for 30sec/68°C for 1min; 30 cycles each) using the primers M5'-3 + M3'-3 (M-1st), M5'-4 + M5'-4 (M-2nd), L8_5'-1 + L8_3'-1 (L8-1st), L8_5'-2 + L8_3'-2 (L8-2nd), R2_5'-1 + R2_3'-1 (R2-1st), and R2_5'-2 + R2_3'-2 (R2-2nd) and analyzed in 1% agarose gel. As a positive control, the constitutively expressed *RPL21* mRNA was detected with RPL21-FW and RPL21-RV.

Cell Fractionation

Cell fractionation was performed as described (Ali *et al*, 2018). 1.4 x 10^6 cells were incubated with 1 mL of Lysis buffer (50 mM Tris pH 8.0, 100 mM NaCl, 5 mM MgCl2, 1 mM EDTA, 0.5% Triton X-100, 1x Complete Protease Inhibitor) on ice for 30 min and spined own at 15,000g for 10 min at 4ºC. The proteins in 500 µL of supernatant were precipitated by adding 10% (final) TCA, washed three times with ice-cold acetone and dissolved in 70 µL 1xSDS sample buffer, while the precipitate was dissolved in 2 mL of fresh Lysis buffer, sonicated (5 times of 20% power-50% pulse-5 pulses with >20 sec interval) using a probe sonicator (Omni-Ruptor 250 Ultrasonic Homogenizer), and then centrifuged at 20,000g for 10 min at 4ºC. Proteins in 1 mL of the supernatant (solubilized chromatin) were precipitated by adding 10% (final) TCA, washed three times with ice-cold acetone and dissolved in 70 µL 1xSDS sample buffer. For cell Fractionation with RNase treatment, 1.4 x 10^6 cells were incubated with 1 mL of Lysis buffer containing 20 µg/mL RNase A on ice for 30 min and spun own at 15,000g for 10 min at 4ºC. The proteins in 1 mL of the supernatant were precipitated by adding 10% (final) TCA, washed three times with ice-cold acetone and dissolved in 70 µL 1xSDS sample buffer, while the precipitate was directly dissolved in 140 µL of 1x SDS sample buffer.

Establishment of *SPT6* germline KO strains

To create the knockout construct for *SPT6*, the 5' and 3' flanking regions of the SPT6 gene were amplified by PCR using the primer sets SPT6-KO-5F-FW/SPT6-KO-5F-RV and SPT6-KO-3F-FW/SPT6-KO-3F-RV, respectively and concatenated by overlap PCR using SPT6-KO-5F-FW and SPT6-KO-3F-RV. The PCR product was cloned into the XhoI and BamHI digested pBlueScriptSK(+). Then the obtained plasmid was digested with BamHI and the pm-resistant cassette *neo4*, which was isolated from the pNeo4 plasmid (Mochizuki, 2008) by SmaI digestion, was inserted. All assembly reactions were done using NEBuilder HiFi DNA assembly mix (NEB). The resulting plasmid, pSPT6-KO-neo4, was digested with SacII and XhoI to release the targeting construct from the vector backbone and introduced into the MIC of mating UMPS214 and UMPS811 cells (Vogt & Mochizuki, 2013) by biolistic

transformation (Cassidy-Hanley et al., 1997). Heterozygous progenies were selected with 100 µg/mL pm and 500 µg/mL 5-Fluoroorotic Acid (5-FOA, Fermentas) in the presence of 1 µg/ml $CdCl_2$. The resulting heterozygous strains were cultured for 10 passages without pm and 5-FOA, and pm-sensitive (pm-s) heterozygous strains were isolated. The pm-s heterozygous cells were then mated with B*VI or B*VII to obtain their round I exconjugants. Finally, their genotypes were examined by PCR to select *SPT6* homozygous heterokaryon KO strains.

Establishment of wild-type and mutant *SPT6* strains

SPT6 cDNA was amplified from total RNA of mating B2086 and CU428 at 6 hpm with SPT6-FW-Turbo and SPT6-RV-Turbo, cloned into SpeI and BamHI sites of pBCMB1-HA-Turbo to obtain pBCMB1-HA-Turbo-SPT6. Then, the *MTT1* promoter and the HA-Turbo-tag were removed using AgeI and XmaI and replaced by the *SPT6* promoter and the N-terminal HA tag, using the overlapping product of two genomic PCR fragments amplified with SPT6-5F-FW, SPT6-5F-RV and with SPT6-N-FW-HA and SPT6-N-RV to obtain pBCSB-HA-SPT6. To produce pBCSB-HA-SPT6-DOL-KR, the SPT6 cDNA lacking the first 248 amino acids were amplified using SPT6_M249_FW and SPT6_M249_RV, replaced with the NdeI-BamHI fragment of pBCSB-HA-SPT6 and then the synthetic DNA SPT6-DOL-KtoR was inserted into the BamHI site. To make pBCSB-HA-SPT6-N-KR, pBCSB-HA-SPT6-M-KR and pBCSB-HA-SPT6-C-KR, the conserved region of *SPT6* was excised from pBCSB-HA-SPT6 using NdeI and AgeI and replaced respectively with Synthetic DNA Set 1 (SPT6-gBlock-118KR-1, SPT6-gBlock-42KR-2, and SPT6-gBlock-42KR-3), Synthetic DNA Set 2 (SPT6-gBlock-42KR-1, SPT6-gBlock-118KR-2, and SPT6-gBlock-42KR-3) or Synthetic DNA Set 3 (SPT6-gBlock-42KR-1, SPT6-gBlock-42KR-2, and SPT6-gBlock-118KR-3). All assembly reactions were done using NEBuilder HiFi DNA assembly mix (NEB). The constructs were digested with XhoI and introduced into the *BTU1* locus of the new MAC of conjugating two *SPT6* KO homozygous heterokaryon strains at 10 hpm by electroporation. Progeny rescued by the introduced constructs were selected in 100 µg/mL paromomycin and 1 µg/mL $CdCl_2$, and cells were phenotypically assorted until cells grew in ~100-120 µg/mL cycloheximide.

Acknowledgements

We acknowledge the MRI facility, member of the national infrastructure France-BioImaging supported by the French National Research Agency (ANR-10-INBS-04), the Proteomics Platform (PPM) of BioCampus Montpellier, and the NGS unit of Vienna BioCenter Core Facilities. This work was supported by an Advanced Grant from the "Investissements d'avenir" program Labex EpiGenMed (ANR-10-LABX-12-01) and an "Accueil de Chercheurs de Haut Niveau" grant (ANR-16-ACHN-0017) from the French National Research Agency to KM, and a "Fin de these" program fellowship from Fondation pour la Recherche Médicale (FRM) to SS.

Author Contributions

SS, TN and KM conceived and planned the experiments. All authors conducted experiments and contributed to the interpretation of the results. SS and KM prepared the manuscript.

Data Deposition

The small RNA sequencing data have been deposited in the Gene Expression Omnibus (GEO) database, https://www.ncbi.nlm.nih.gov/geo (accession no. GSE243435). The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (https://www.ebi.ac.uk/pride/) with reference 1-20230915-080058-2260772.

References

Ali EI, Loidl J & Howard-Till RA (2018) A streamlined cohesin apparatus is sufficient for mitosis and meiosis in the protist Tetrahymena. *Chromosoma* 127: 421–435

Andersen PR, Tirian L, Vunjak M & Brennecke J (2017) A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature* 549: 54–59

Andreev VI, Yu C, Wang J, Schnabl J, Tirian L, Gehre M, Handler D, Duchek P, Novatchkova M, Baumgartner L, *et al* (2022) Panoramix SUMOylation on chromatin connects the piRNA pathway to the cellular heterochromatin machinery. *Nat Struct Mol Biol* 29: 130–142

Aronica L, Bednenko J, Noto T, DeSouza LV, Siu KWM, 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

Austerberry CF, Allis CD & Yao MC (1984) Specific DNA rearrangements in synchronously developing nuclei of Tetrahymena. *Proc Natl Acad Sci USA* 81: 7383–7387

Beauclair G, Bridier-Nahmias A, Zagury J-F, Saïb A & Zamborlini A (2015) JASSA: a comprehensive tool for prediction of SUMOylation sites and SIMs. *Bioinformatics* 31: 3483–3491

Bruns PJ & Cassidy-Hanley D (2000) Biolistic transformation of macro- and micronuclei. *Methods Cell Biol* 62: 501–512

Chalker DL, Meyer E & Mochizuki K (2013) Epigenetics of ciliates. *Cold Spring Harb Perspect Biol* 5: a017764

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

Cheng C-Y, Vogt A, Mochizuki K & Yao M-C (2010) A domesticated piggyBac transposase plays key roles in heterochromatin dynamics and DNA cleavage during programmed DNA deletion in Tetrahymena thermophila. *Mol Biol Cell* 21: 1753–1762

Chen ES, Zhang K, Nicolas E, Cam HP, Zofall M & Grewal SIS (2008) Cell cycle control of centromeric repeat transcription and heterochromatin assembly. *Nature* 451: 734–737

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

Duan X, Sarangi P, Liu X, Rangi GK, Zhao X & Ye H (2009) Structural and functional insights into the roles of the Mms21 subunit of the Smc5/6 complex. *Mol Cell* 35: 657–668

Garg J, Saettone A, Nabeel-Shah S, Cadorin M, Ponce M, Marquez S, Pu S, Greenblatt J, Lambert J-P, Pearlman RE, *et al* (2019) The Med31 Conserved Component of the Divergent Mediator Complex in Tetrahymena thermophila Participates in Developmental Regulation. *Curr Biol* 29: 2371-2379.e6

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

Hamilton EP & Orias E (2000) Genetic crosses: setting up crosses, testing progeny, and isolating phenotypic assortants. *Methods Cell Biol* 62: 219–228

Hendriks IA & Vertegaal ACO (2016) Label-Free Identification and Quantification of SUMO Target Proteins. *Methods Mol Biol* 1475: 171–193

Hochstrasser M (2001) SP-RING for SUMO: new functions bloom for a ubiquitin-like protein. *Cell* 107: 5–8

Holoch D & Moazed D (2015) RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet* 16: 71–84

Kataoka K & Mochizuki K (2015) Phosphorylation of an HP1-like Protein Regulates Heterochromatin Body Assembly for DNA Elimination. *Dev Cell* 35: 775–788
Kataoka K & Mochizuki K (2017) Heterochromatin aggregation during DNA elimination in Tetrahymena is facilitated by a prion-like protein. *J Cell Sci* 130: 480–489
Kataoka K, Schoeberl UE & Mochizuki K (2010) Modules for C-terminal epitope tagging of Tetrahymena genes. *J Microbiol Methods* 82: 342–346
Kim H, Ding Y-H, Zhang G, Yan Y-H, Conte D, Dong M-Q & Mello CC (2021) HDAC1 SUMOylation promotes Argonaute-directed transcriptional silencing in C. elegans. *eLife* 10
Kloc A, Zaratiegui M, Nora E & Martienssen R (2008) RNA interference guides histone modification during the S phase of chromosomal replication. *Curr Biol* 18: 490–495
Law JA, Du J, Hale CJ, Feng S, Krajewski K, Palanca AMS, Strahl BD, Patel DJ & Jacobsen SE (2013) Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* 498: 385–389
Law JA, Vashisht AA, Wohlschlegel JA & Jacobsen SE (2011) SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. *PLoS Genet* 7: e1002195
Lepère G, Bétermier M, Meyer E & Duharcourt S (2008) Maternal noncoding transcripts antagonize the targeting of DNA elimination by scanRNAs in Paramecium tetraurelia. *Genes Dev* 22: 1501–1512
Liebelt F, Jansen NS, Kumar S, Gracheva E, Claessens LA, Verlaan-de Vries M, Willemstein E & Vertegaal ACO (2019) The poly-SUMO2/3 protease SENP6 enables assembly of the constitutive centromere-associated network by group deSUMOylation. *Nat Commun* 10: 3987
Liu Y, Taverna SD, Muratore TL, Shabanowitz J, Hunt DF & Allis CD (2007) RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in Tetrahymena. *Genes Dev* 21: 1530–1545
Loidl J & Scherthan H (2004) Organization and pairing of meiotic chromosomes in the ciliate Tetrahymena thermophila. *J Cell Sci* 117: 5791–5801
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
Matsuda A & Forney JD (2006) The SUMO pathway is developmentally regulated and required for programmed DNA elimination in Paramecium tetraurelia. *Eukaryotic Cell* 5: 806–815
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
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) RNA polymerase II localizes in Tetrahymena thermophila meiotic micronuclei when micronuclear transcription associated with genome rearrangement occurs. *Eukaryotic Cell* 3: 1233–1240
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

Mutazono M, Noto T & Mochizuki K (2019) Diversification of small RNA amplification mechanisms for targeting transposon-related sequences in ciliates. *Proc Natl Acad Sci USA* 116: 14639–14644

Nasir AM, Yang Q, Chalker DL & Forney JD (2015) SUMOylation is developmentally regulated and required for cell pairing during conjugation in Tetrahymena thermophila. *Eukaryotic Cell* 14: 170–181

Ninova M, Chen Y-CA, Godneeva B, Rogers AK, Luo Y, Fejes Tóth K & Aravin AA (2020a) Su(var)2-10 and the SUMO Pathway Link piRNA-Guided Target Recognition to Chromatin Silencing. *Mol Cell* 77: 556-570.e6

Ninova M, Godneeva B, Chen Y-CA, Luo Y, Prakash SJ, Jankovics F, Erdélyi M, Aravin AA & Fejes Tóth K (2020b) The SUMO Ligase Su(var)2-10 Controls Hetero- and Euchromatic Gene Expression via Establishing H3K9 Trimethylation and Negative Feedback Regulation. *Mol Cell* 77: 571-585.e4

Noto T, Kataoka K, Suhren JH, Hayashi A, Woolcock KJ, Gorovsky MA & Mochizuki K (2015) Small-RNA-Mediated Genome-wide trans-Recognition Network in Tetrahymena DNA Elimination. *Mol Cell* 59: 229–242

Noto T, Kurth HM, Kataoka K, Aronica L, DeSouza LV, Siu KWM, Pearlman RE, Gorovsky MA & Mochizuki K (2010) The Tetrahymena argonaute-binding protein Giw1p directs a mature argonaute-siRNA complex to the nucleus. *Cell* 140: 692–703

Noto T & Mochizuki K (2017) Whats, hows and whys of programmed DNA elimination in *Tetrahymena*. *Open Biol* 7: 170172

Noto T & Mochizuki K (2018) Small RNA-mediated trans-nuclear and trans-element communications in *Tetrahymena* DNA elimination. *Curr Biol* 28: 1938-1949.e5

Praefcke GJK, Hofmann K & Dohmen RJ (2012) SUMO playing tag with ubiquitin. *Trends Biochem Sci* 37: 23–31

Schoeberl UE, Kurth HM, Noto T & Mochizuki K (2012) Biased transcription and selective degradation of small RNAs shape the pattern of DNA elimination in *Tetrahymena. Genes Dev* 26: 1729–1742

Schoeberl UE & Mochizuki K (2011) Keeping the soma free of transposons: programmed DNA elimination in ciliates. *J Biol Chem* 286: 37045–37052

Schönborn J, Oberstrass J, Breyel E, Tittgen J, Schumacher J & Lukacs N (1991) Monoclonal antibodies to double-stranded RNA as probes of RNA structure in crude nucleic acid extracts. *Nucleic Acids Res* 19: 2993–3000

Shehzada S & Mochizuki K (2022) PIWI-Directed DNA Elimination for Tetrahymena Genetics. *Methods Mol Biol* 2509: 53–68

Shin JA, Choi ES, Kim HS, Ho JCY, Watts FZ, Park SD & Jang YK (2005) SUMO modification is involved in the maintenance of heterochromatin stability in fission yeast. *Mol Cell* 19: 817–828

Staudinger JL (2017) The molecular interface between the SUMO and ubiquitin systems. *Adv Exp Med Biol* 963: 99–110

Tian M, Mochizuki K & Loidl J (2019) Non-coding RNA Transcription in Tetrahymena Meiotic Nuclei Requires Dedicated Mediator Complex-Associated Proteins. *Curr Biol* 29: 2359-2370.e5

Vogt A & Mochizuki K (2013) A domesticated PiggyBac transposase interacts with heterochromatin and catalyzes reproducible DNA elimination in *Tetrahymena*. *PLoS Genet* 9: e1004032

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

Wilson RC & Doudna JA (2013) Molecular mechanisms of RNA interference. *Annu Rev Biophys* 42: 217–239

Woo T-T, Chao J-L & Yao M-C (2016) Dynamic distributions of long double-stranded RNA in Tetrahymena during nuclear development and genome rearrangements. *J Cell Sci* 129: 1046–1058

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

Xiong J, Lu X, Zhou Z, Chang Y, Yuan D, Tian M, Zhou Z, Wang L, Fu C, Orias E, *et al* (2012) Transcriptome analysis of the model protozoan, Tetrahymena thermophila, using Deep RNA sequencing. *PLoS ONE* 7: e30630

Xu J, Zhao X, Mao F, Basrur V, Ueberheide B, Chait BT, Allis CD, Taverna SD, Gao S, Wang W, *et al* (2021) A Polycomb repressive complex is required for RNAi-mediated heterochromatin formation and dynamic distribution of nuclear bodies. *Nucleic Acids Res* 49: 5407–5425

Yao M-C, Yao C-H, 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

Yunus AA & Lima CD (2009) Structure of the Siz/PIAS SUMO E3 ligase Siz1 and determinants required for SUMO modification of PCNA. *Mol Cell* 35: 669–682

Figure Legends

Figure 1. Ema2 is expressed during conjugation and localized in the MAC

(A) The expression profile of *EMA2* mRNA. On the X-axis, L1 to Lh represent growing cells, and S0 to S24 represent the starvation of *Tetrahymena* cells from 0 to 24 hours. C0 to C18 represent cells in the conjugation and postconjugation stages from 0 to 18 hpm. On the Y-axis, the number represents the relative expression of *EMA2* mRNA. The mRNA expression data were obtained from (Miao *et al*, 2009).

(B) Ema2 localization. Two Ema2-HA strains were mated and fixed at the indicated time points (Vg = vegetative cell). An anti-HA antibody was used to localize Ema2-HA (green), and DNA was stained with DAPI (magenta). The MIC (i), the parental MAC (a) and the newly formed MAC (na) are marked with arrowheads. All pictures share the scale bar.

Figure 2. Ema2 is required for completing DNA elimination

(A) Two wild-type (WT) or two *EMA2 KO* (KO) cell lines were mixed, and their exconjugants at 36 hpm were analyzed by DNA-FISH with florescent probes complementary to the Tlr1 element (green). DNA was counterstained with DAPI (magenta). The MICs (i) and the new MACs (na) are marked with arrowheads. All pictures share the scale bar.

(B) Exconjugants from wild-type (WT) cells, *EMA2* KO cells, and *TWI1* KO cells were stained as in (A), and the IES retention index was calculated (see Materials and Methods for details).

(C) Two wild-type (WT) cells and two independent crosses of *EMA2* KO cells (Cross 1 and 2) were mated, and the conjugating pairs were isolated for the viability test. The percentages of pairs that gave rise to viable sexual progeny are shown. “n” represents the number of total pairs tested.

Figure 3. Ema2 is required for TDSD

(A) Total RNA was isolated from conjugating wild-type (WT) and *EMA2* KO (KO) cells at 3, 4.5 and 6 hpm, separated in denaturing gel and stained with the nucleic acid dye Gel-Red (right). Then, RNA was transferred to a membrane and hybridized with the radioactive Mi-9 probe, which is complementary to a repetitive sequence in MDS (right).

(B) Small RNAs from conjugation wild-type (WT), *EMA1* KO, and two independent crosses of *EMA2* KO strains were isolated at 3 and 8 hpm and analyzed by high-throughput sequencing. Normalized numbers (RPKM [read per kilobase of unique sequence per million]) of sequenced small RNAs (26-32 nt) that uniquely matched to the MDS (left) or IES (right) genomic tiles

(see Materials and Methods) are shown as box plots. The median value is represented by the horizontal bar in the box. The minimum and maximum values are indicated by the bars on top and bottom of the box, respectively, with 1.5x IQR.

Figure 4. Ema2 acts as a SUMO E3 ligase

(A) The SP-RING domain of Ema2 is compared with that of MMS21 and SIZ1 in *S. cerevisiae*, Su(var)2-10 in *D. melanogaster* and PIAS1 in *H. sapiens*. The conserved cysteine and histidine residues that are involved in zinc ion binding are highlighted in yellow. The residues that stabilize the domain structure of some SP-RING domain proteins are marked with pink.

(B) GST alone (GST), GST-tagged Ema2 (GST-Ema2) and His-tagged Ubc9 (His-Ubc9) were recombinantly expressed in *E. coli* and purified. GST and GST-Ema2 were immobilized on glutathione beads and incubated with His-Ubc9. Proteins retained on the beads were eluted, and the input and eluted proteins (PD) were analyzed by western blotting using anti-GST (left) and anti-His (right) antibodies.

(C) An *EMA2* KO strain expressing HA-tagged Smt3 was crossed with a wild-type strain (WT cross, WT) or another *EMA2* KO strain (*EMA2* KO cross, KO), and their total proteins at the indicated time points were analyzed by western blotting using anti-HA (top), anti-Twi1 (middle) and anti-alpha tubulin (bottom) antibodies.

Figure 5. Ema2 promotes SUMOylation of Spt6.

(A) A construct expressing His-tagged Smt3 (His-Smt3) was introduced into a wild-type and one of the *EMA2* KO strains and crossed with another wild-type (WT cross) or one of the other *EMA2* KO strains (*EMA2* KO cross). Proteins were harvested at 6 hpm, and His-Smt3-conjugated proteins were purified with Ni-NTA beads from conjugating and analyzed by mass spectrometry. Values of log base 2 of label-free quantification (LFQ) intensities of each identified protein between the WT cross (WT) and *EMA2* KO cross (*EMA2* KO) were compared.

(B) A construct expressing HA-tagged Spt6 (Spt6-HA) was introduced into an *EMA2* KO strain and crossed with a wild-type (WT-cross, WT) or another *EMA2* KO (KO-cross, KO) strain. Total proteins were harvested at 4.5 and 6 hpm, and Spt6-HA was detected by western blotting using an anti-HA antibody. Twi1 and alpha tubulin (Tub) were also analyzed to monitor mating efficiency and loading, respectively. The positions of modified and unmodified Spt6-HA proteins are marked with a bracket and an arrowhead, respectively.

(C) Total proteins were harvested from WT and KO crosses at 6 hpm (input), and Spt6-HA was immunoprecipitated using an anti-HA antibody (IP). The purified proteins were analyzed by western blotting using an anti-HA (left) or an anti-Smt3 (right) antibody. SUMOylated Spt6 proteins are marked with brackets. Unidentified protein cross reacting with the anti-HA antibody is marked with an asterisk.

Figure 6. Ema2 is required for the accumulation of lncRNA transcripts from the parental MAC.

(A) (Left) Schematic representation of the RT–PCR assay. The black bars and the open box represent the MDS and IES, respectively. The arrows represent the primers used for RT–PCR. The lengths of the PCR amplicons are shown with double-sided arrows. (Right) Wild-type (WT) or *EMA2 KO* (KO) cells were mated, and their total RNAs at 6 hpm were used for RT–PCR. The positions corresponding to the PCR products of lncRNAs from the MAC M, L8, and R2 loci are marked. *RPL21* mRNA was also analyzed as a positive control.

(B) Conjugating wild-type (WT), *EMA1* KO and *EMA2* KO cells at the indicated time points were analyzed by immunofluorescence staining using the anti-long dsRNA antibody J2 (dsRNA, green). DNA was counterstained with DAPI (magenta). The MIC, the parental MAC and the new MAC are indicated with single arrowhead, double arrowhead, and single arrowhead with “n”, respectively. All pictures share the scale bar.

(C) Heterochromatin formation in the parental MAC. Conjugating wild-type (WT), *EMA2* KO, *EMA1* KO, *TWI1* KO, and *EZL1* KO cells at 6 hpm were analyzed by immunofluorescence staining using an anti-H3K27me3 antibody (green). DNA was counterstained with DAPI (magenta). The MIC and the parental MAC are indicated with single and double arrowheads, respectively. All pictures share the scale bar.

Figure 7. Ema2 facilitates the chromatin association of Spt6 and RNAPII

(A) A construct expressing HA-tagged Spt6 (Spt6-HA) was introduced into an *EMA2* KO strain and crossed with a wild-type (WT-cross, WT) or another *EMA2* KO (KO-cross, KO) strain. The localizations of Spt6-HA and Rpb3, the third largest subunit of RNA polymerase II (RNAPII), were analyzed by immunofluorescence staining at 4.5 hpm using anti-HA (left) or anti-Rpb3 (right) antibodies, respectively. DNA was counterstained with DAPI (magenta). The parental MACs are marked with arrowheads. All other structures stained by DAPI are MICs. All pictures share the scale bar.

(B) (Left) Schematic representation of the cell fractionation assay. Conjugating cells were incubated with a lysis buffer that releases cytoplasmic (green) and nucleoplasmic (magenta) proteins to the soluble fraction (S1). Then, the insoluble fraction was resuspended in fresh lysis buffer and sonicated. The solubilized fraction after sonication (S2) contains fragmented chromatin. (Right) S1 and S2 fractions as well as total cellular proteins (W) from WT-cross (WT) and *EMA2* KO-cross (KO) explained in (A) at 4.5 hpm were analyzed by western blotting using anti-HA (Spt6-HA), anti-Rpb3 and anti-histone H3 (bottom) antibodies. The position of SUMOylated Spt6-HA is marked with a bracket.

(C) (Left) Schematic representation of the cell fractionation assay with RNase treatment. Conjugating cells were incubated with a lysis buffer that releases cytoplasmic (green) and nucleoplasmic (magenta) proteins to the soluble fraction (S1). Then, the insoluble fraction was incubated in fresh lysis buffer with (+) or without (-) RNase A, and the solubilized (S2) and insoluble (I) fractions were obtained. (Right) S1, S2 and I fractions as well as total cellular proteins (W) from *SPT6* KO strains rescued with the *HA-SPT6-WT* construct (see Figure 8) were analyzed by western blotting using anti-HA (HA-Spt6), anti-Twi1p, anti-Rpb3 and anti-histone H3 (bottom) antibodies. The position of SUMOylated HA-Spt6 is marked with a bracket.

Figure 8. Spt6 SUMOylation is dispensable for parental MAC lncRNA transcription

(A) Schematic representation of *SPT6* mutant constructs with lysine (K)-to-arginine (R) substitutions. Spt6 has a domain of low complexity (DOL) region followed by conserved TEX-like and SH2 domains. The DOL-KR mutant has K to R substitutions for all lysine residues in the DOL region. N-KR, M-KR and C-KR mutants have K to R substitutions for lysine residues that were on the surface of a predicted Spt6 structure in one of the three non-DOL regions (N, M and C-regions) in addition to those surface lysine residues that were also predicted to be SUMOylatable by an algorithm in all three regions.

(B) Total proteins were harvested at 4.5 hpm from conjugating *SPT6* KO cells rescued with wild-type (*SPT6-WT*), *SPT6-DOL-KR* or *SPT6-C-KR* constructs. Because mating was low in the SPT6-C-KR rescued cells, 3 times more total protein sample was loaded for these cells. Spt6-HA was detected by western blotting using an anti-HA antibody. The positions of modified and unmodified Spt6-HA proteins are marked with “M” and “UM”, respectively.

(C) *SPT6-WT* (*WT*) and *SPT6-C-KR* (*C-KR*) rescued cells at 4.5 hpm were analyzed by cell fractionation as described in Figure 7B. The position of SUMOylated Spt6-HA is marked with a bracket.

(D) Accumulation of long dsRNAs in *SPT6-WT* (*WT*) and *SPT6-C-KR* (*C-KR*) rescued cells at 4.5 hpm was analyzed by immunofluorescence staining using the J2 antibody as described in Figure 6B. The parental MACs are marked with arrowheads. All other structures stained by DAPI are MICs. All pictures share the scale bar.

(E) *SPT6-WT* (*WT*) and *SPT6-C-KR* (*C-KR*) rescued cells at 36 hpm were analyzed by DNA-FISH with florescent probes complementary to the Tlr1 element (red). DNA was counterstained with DAPI (blue). The MICs (i) and the new MACs (na) are marked with arrowheads. All pictures share the scale bar.

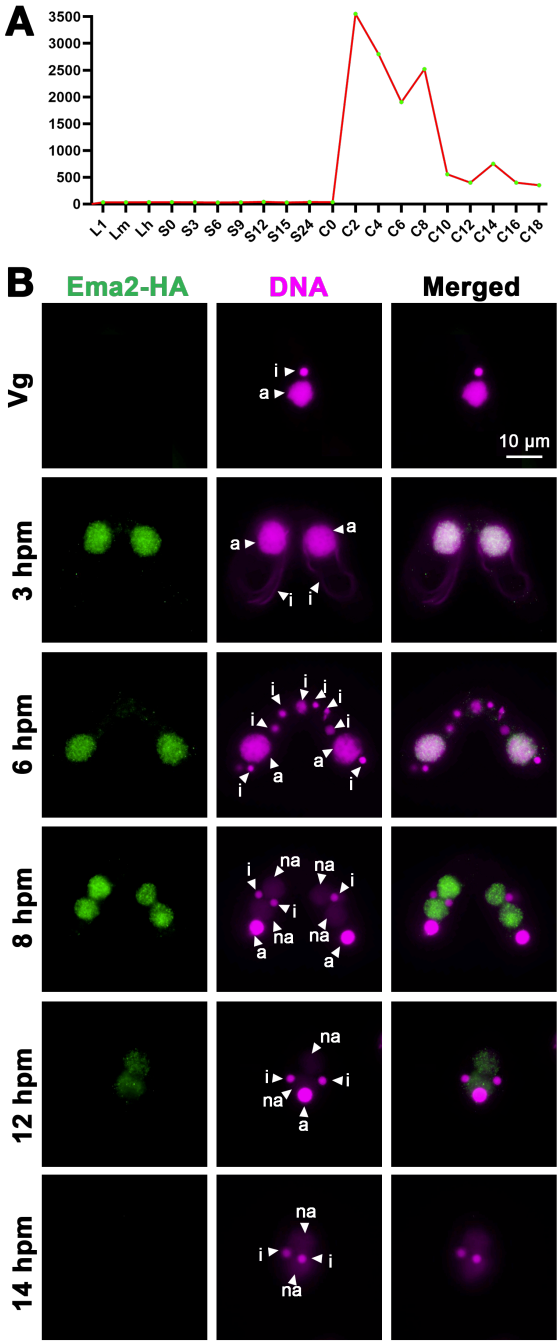
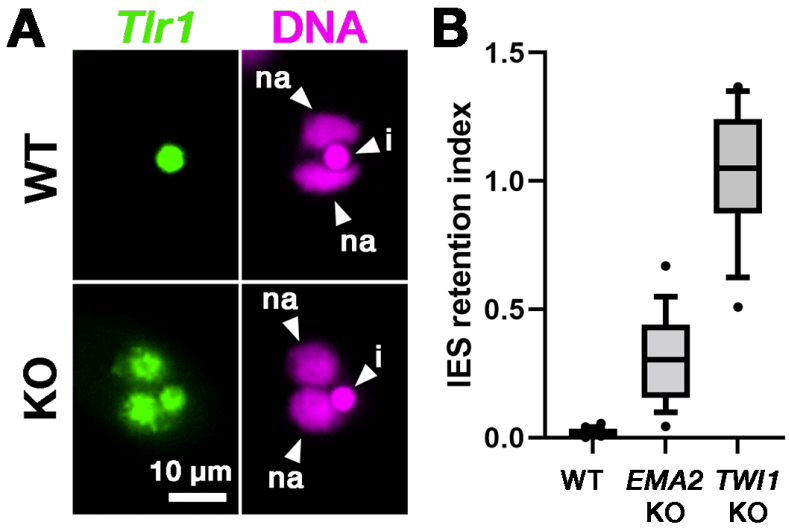


Figure 1. Ema2 is expressed during conjugation and localized in the MAC

(A) The expression profile of *EMA2* mRNA. On the X-axis, L1 to Lh represent growing cells, and S0 to S24 represent the starvation of *Tetrahymena* cells from 0 to 24 hours. C0 to C18 represent cells in the conjugation and postconjugation stages from 0 to 18 hpm. On the Y-axis, the number represents the relative expression of *EMA2* mRNA. The mRNA expression data were obtained from (Miao *et al*, 2009).

(B) Ema2 localization. Two Ema2-HA strains were mated and fixed at the indicated time points (Vg = vegetative cell). An anti-HA antibody was used to localize Ema2-HA (green), and DNA was stained with DAPI (magenta). The MIC (i), the parental MAC (a) and the newly formed MAC (na) are marked with arrowheads. All pictures share the scale bar.



C

Type of Mating Cells	Percentage of Pairs Producing Viable Progeny
WT	76.04 (n=96)
EMA2 KO Cross 1	0 (n=96)
EMA2 KO Cross 2	0 (n=96)

Figure 2. Ema2 is required for completing DNA elimination

(A) Two wild-type (WT) or two *EMA2 KO* (KO) cell lines were mixed, and their exconjugants at 36 hpm were analyzed by DNA-FISH with florescent probes complementary to the Tlr1 element (green). DNA was counterstained with DAPI (magenta). The MICs (i) and the new MACs (na) are marked with arrowheads. All pictures share the scale bar.

(B) Exconjugants from wild-type (WT) cells, *EMA2* KO cells, and *TWI1* KO cells were stained as in (A), and the IES retention index was calculated (see Materials and Methods for details).

(C) Two wild-type (WT) cells and two independent crosses of *EMA2* KO cells (Cross 1 and 2) were mated, and the conjugating pairs were isolated for the viability test. The percentages of pairs that gave rise to viable sexual progeny are shown. “n” represents the number of total pairs tested.

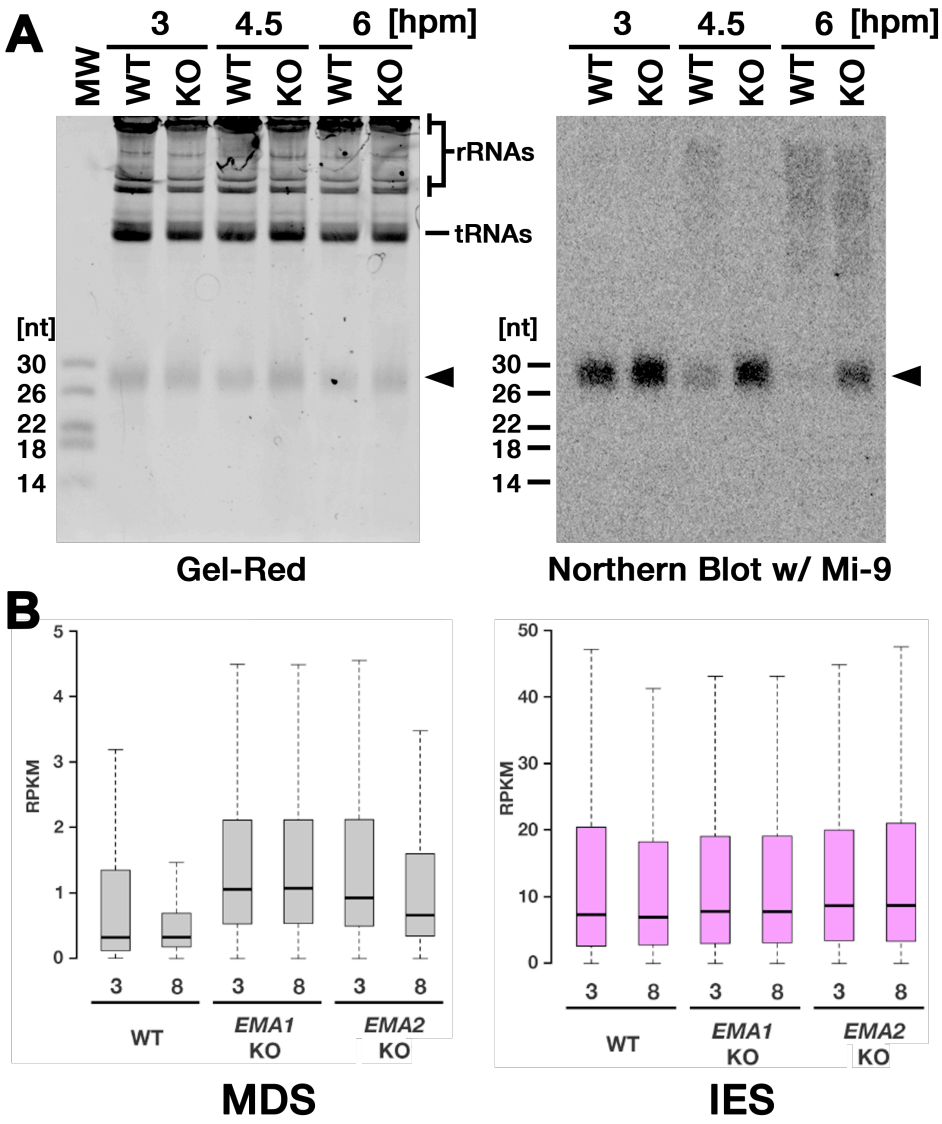


Figure 3. Ema2 is required for TDSD

(A) Total RNA was isolated from conjugating wild-type (WT) and *EMA2* KO (KO) cells at 3, 4.5 and 6 hpm, separated in denaturing gel and stained with the nucleic acid dye Gel-Red (right). Then, RNA was transferred to a membrane and hybridized with the radioactive Mi-9 probe, which is complementary to a repetitive sequence in MDS (right).

(B) Small RNAs from conjugation wild-type (WT), *EMA1* KO, and two independent crosses of *EMA2* KO strains were isolated at 3 and 8 hpm and analyzed by high-throughput sequencing. Normalized numbers (RPKM [read per kilobase of unique sequence per million]) of sequenced small RNAs (26-32 nt) that uniquely matched to the MDS (left) or IES (right) genomic tiles (see Materials and Methods) are shown as box plots. The median value is represented by the horizontal bar in the box. The minimum and maximum values are indicated by the bars on top and bottom of the box, respectively, with 1.5x IQR.

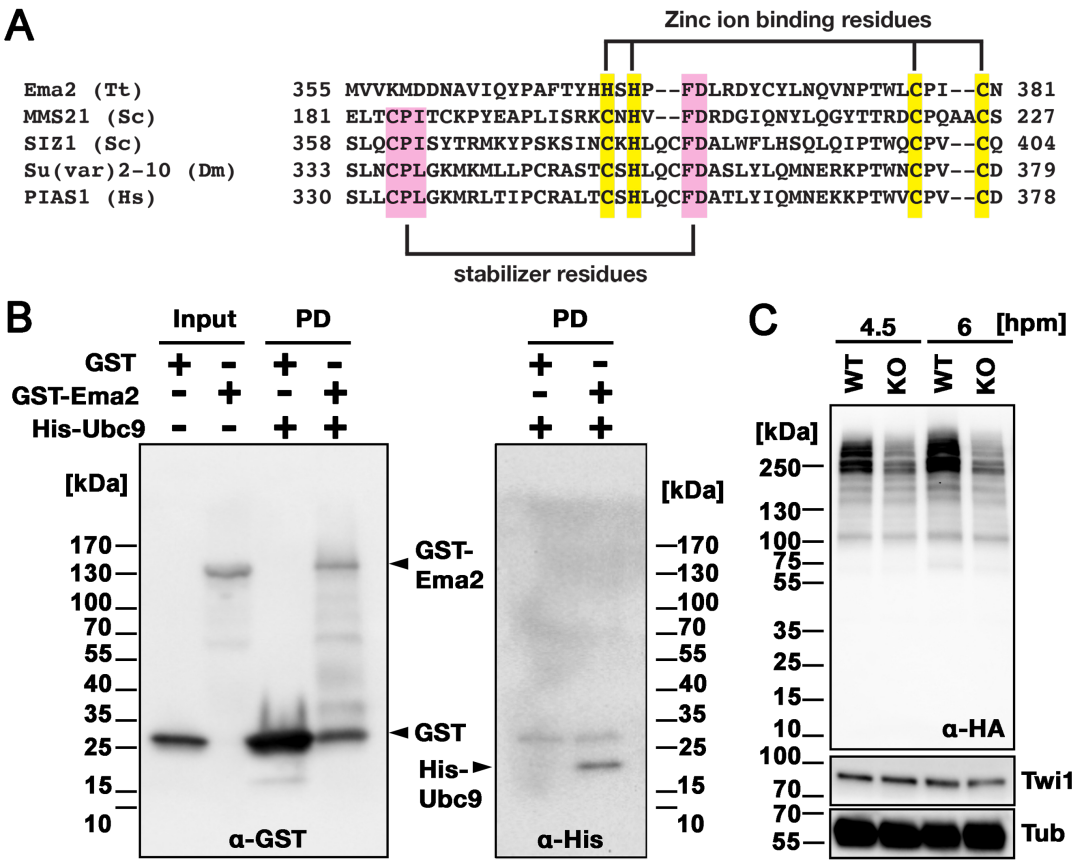


Figure 4. Ema2 acts as a SUMO E3 ligase

(A) The SP-RING domain of Ema2 is compared with that of MMS21 and SIZ1 in *S. cerevisiae*, Su(var)2-10 in *D. melanogaster* and PIAS1 in *H. sapiens*. The conserved cysteine and histidine residues that are involved in zinc ion binding are highlighted in yellow. The residues that stabilize the domain structure of some SP-RING domain proteins are marked with pink.

(B) GST alone (GST), GST-tagged Ema2 (GST-Ema2) and His-tagged Ubc9 (His-Ubc9) were recombinantly expressed in *E. coli* and purified. GST and GST-Ema2 were immobilized on glutathione beads and incubated with His-Ubc9. Proteins retained on the beads were eluted, and the input and eluted proteins (PD) were analyzed by western blotting using anti-GST (left) and anti-His (right) antibodies.

(C) An *EMA2* KO strain expressing HA-tagged Smt3 was crossed with a wild-type strain (WT cross, WT) or another *EMA2* KO strain (*EMA2* KO cross, KO), and their total proteins at the indicated time points were analyzed by western blotting using anti-HA (top), anti-Twi1 (middle) and anti-alpha tubulin (bottom) antibodies.

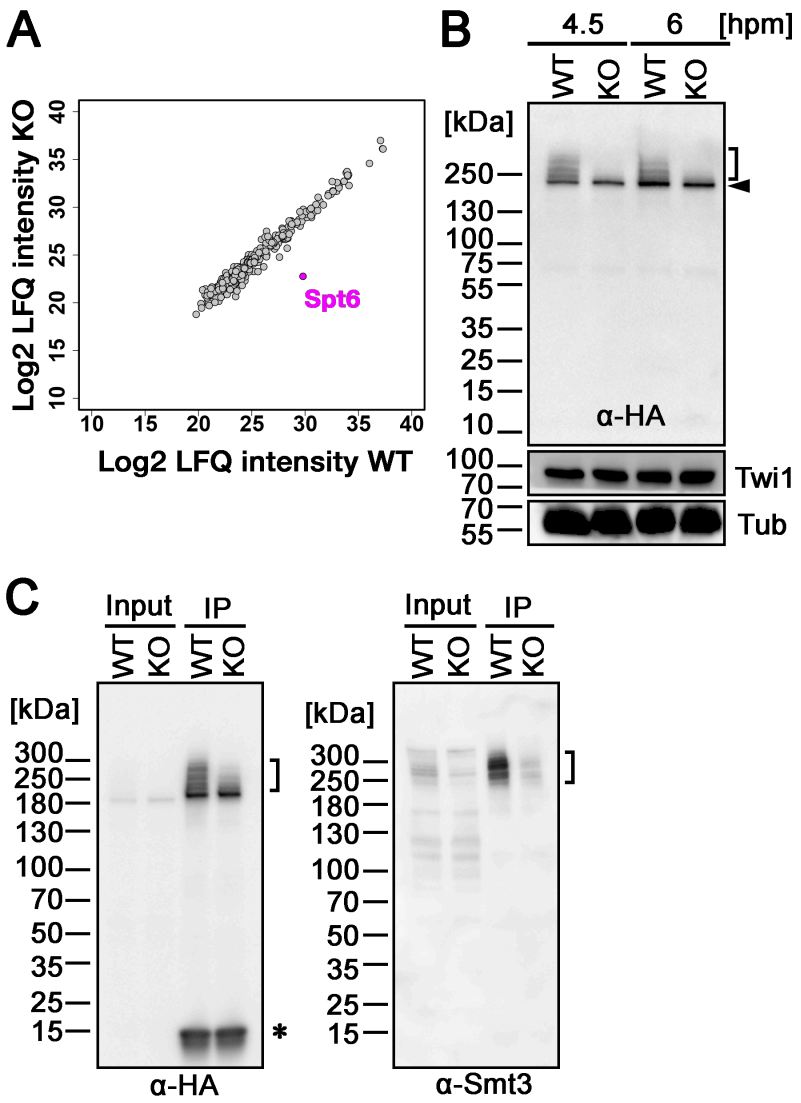


Figure 5. Ema2 promotes SUMOylation of Spt6.

(A) A construct expressing His-tagged Smt3 (His-Smt3) was introduced into a wild-type and one of the *EMA2* KO strains and crossed with another wild-type (WT cross) or one of the other *EMA2* KO strains (*EMA2* KO cross). Proteins were harvested at 6 hpm, and His-Smt3-conjugated proteins were purified with Ni-NTA beads from conjugating and analyzed by mass spectrometry. Values of log base 2 of label-free quantification (LFQ) intensities of each identified protein between the WT cross (WT) and *EMA2* KO cross (*EMA2* KO) were compared.

(B) A construct expressing HA-tagged Spt6 (Spt6-HA) was introduced into an *EMA2* KO strain and crossed with a wild-type (WT-cross, WT) or another *EMA2* KO (KO-cross, KO) strain. Total proteins were harvested at 4.5 and 6 hpm, and Spt6-HA was detected by western blotting using an anti-HA antibody. Twi1 and alpha tubulin (Tub) were also analyzed to monitor mating efficiency and loading, respectively. The positions of modified and unmodified Spt6-HA proteins are marked with a bracket and an arrowhead, respectively.

(C) Total proteins were harvested from WT and KO crosses at 6 hpm (input), and Spt6-HA was immunoprecipitated using an anti-HA antibody (IP). The purified proteins were analyzed by western blotting using an anti-HA (left) or an anti-Smt3 (right) antibody. SUMOylated Spt6 proteins are marked with brackets. Unidentified protein cross reacting with the anti-HA antibody is marked with an asterisk.

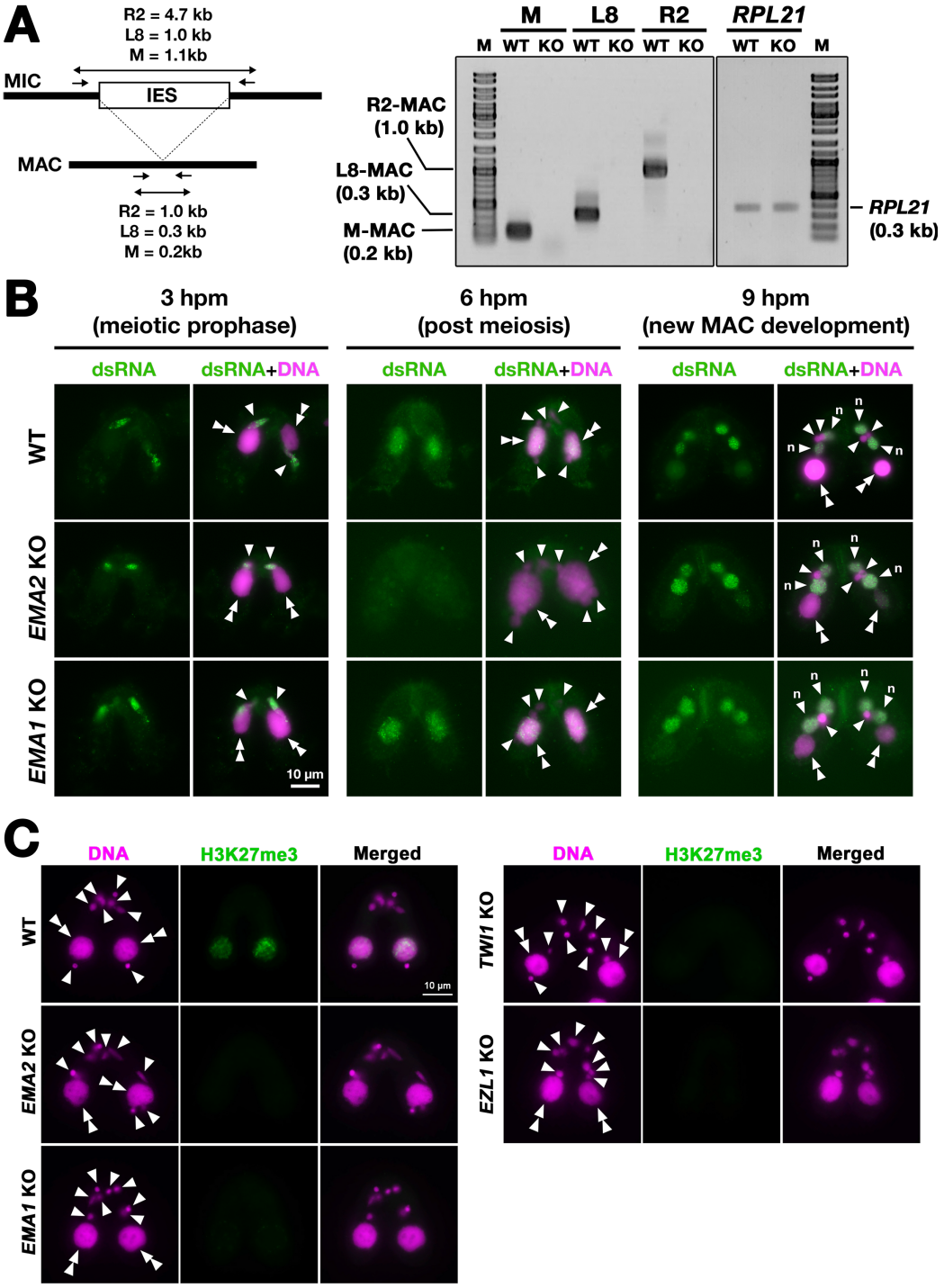


Figure 6. Ema2 is required for the accumulation of lncRNA transcripts from the parental MAC.

(A) (Left) Schematic representation of the RT–PCR assay. The black bars and the open box represent the MDS and IES, respectively. The arrows represent the primers used for RT–PCR. The lengths of the PCR amplicons are shown with double-sided arrows. (Right) Wild-type (WT) or *EMA2 KO* (KO) cells were mated, and their total RNAs at 6 hpm were used for RT–PCR. The positions corresponding to the PCR products of lncRNAs from the MAC M, L8, and R2 loci are marked. *RPL21* mRNA was also analyzed as a positive control.

(B) Conjugating wild-type (WT), *EMA1* KO and *EMA2* KO cells at the indicated time points were analyzed by immunofluorescence staining using the anti-long dsRNA antibody J2 (dsRNA, green). DNA was counterstained with DAPI (magenta). The MIC, the parental MAC and the new MAC are indicated with single arrowhead, double arrowhead, and single arrowhead with "n", respectively. All pictures share the scale bar.

(C) Heterochromatin formation in the parental MAC. Conjugating wild-type (WT), *EMA2* KO, *EMA1* KO, *TWI1* KO, and *EZL1* KO cells at 6 hpm were analyzed by immunofluorescence staining using an anti-H3K27me3 antibody (green). DNA was counterstained with DAPI (magenta). The MIC and the parental MAC are indicated with single and double arrowheads, respectively. All pictures share the scale bar.

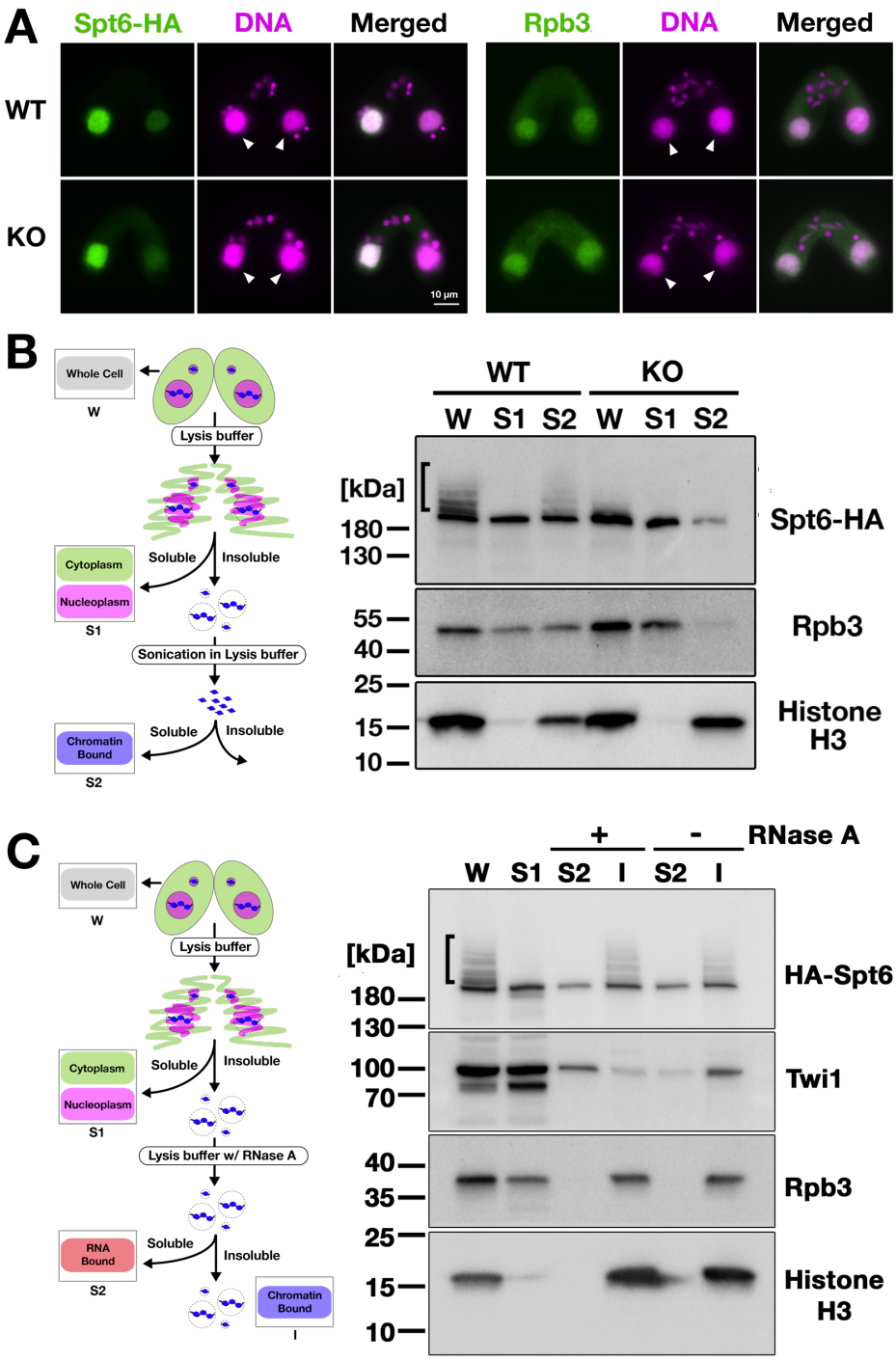


Figure 7. Ema2 facilitates the chromatin association of Spt6 and RNAPII

(A) A construct expressing HA-tagged Spt6 (Spt6-HA) was introduced into an *EMA2* KO strain and crossed with a wild-type (WT-cross, WT) or another *EMA2* KO (KO-cross, KO) strain. The localizations of Spt6-HA and Rpb3, the third largest subunit of RNA polymerase II (RNAPII), were analyzed by immunofluorescence staining at 4.5 hpm using anti-HA (left) or anti-Rpb3 (right) antibodies, respectively. DNA was counterstained with DAPI (magenta). The parental MACs are marked with arrowheads. All other structures stained by DAPI are MICs. All pictures share the scale bar.

(B) (Left) Schematic representation of the cell fractionation assay. Conjugating cells were incubated with a lysis buffer that releases cytoplasmic (green) and nucleoplasmic (magenta) proteins to the soluble fraction (S1). Then,

the insoluble fraction was resuspended in fresh lysis buffer and sonicated. The solubilized fraction after sonication (S2) contains fragmented chromatin. (Right) S1 and S2 fractions as well as total cellular proteins (W) from WT-cross (WT) and *EMA2* KO-cross (KO) explained in (A) at 4.5 hpm were analyzed by western blotting using anti-HA (Spt6-HA), anti-Rpb3 and anti-histone H3 (bottom) antibodies. The position of SUMOylated Spt6-HA is marked with a bracket.

(C) (Left) Schematic representation of the cell fractionation assay with RNase treatment. Conjugating cells were incubated with a lysis buffer that releases cytoplasmic (green) and nucleoplasmic (magenta) proteins to the soluble fraction (S1). Then, the insoluble fraction was incubated in fresh lysis buffer with (+) or without (-) RNase A, and the solubilized (S2) and insoluble (I) fractions were obtained. (Right) S1, S2 and I fractions as well as total cellular proteins (W) from *SPT6* KO strains rescued with the *HA-SPT6-WT* construct (see Figure 8) were analyzed by western blotting using anti-HA (HA-Spt6), anti-Twi1p, anti-Rpb3 and anti-histone H3 (bottom) antibodies. The position of SUMOylated HA-Spt6 is marked with a bracket.

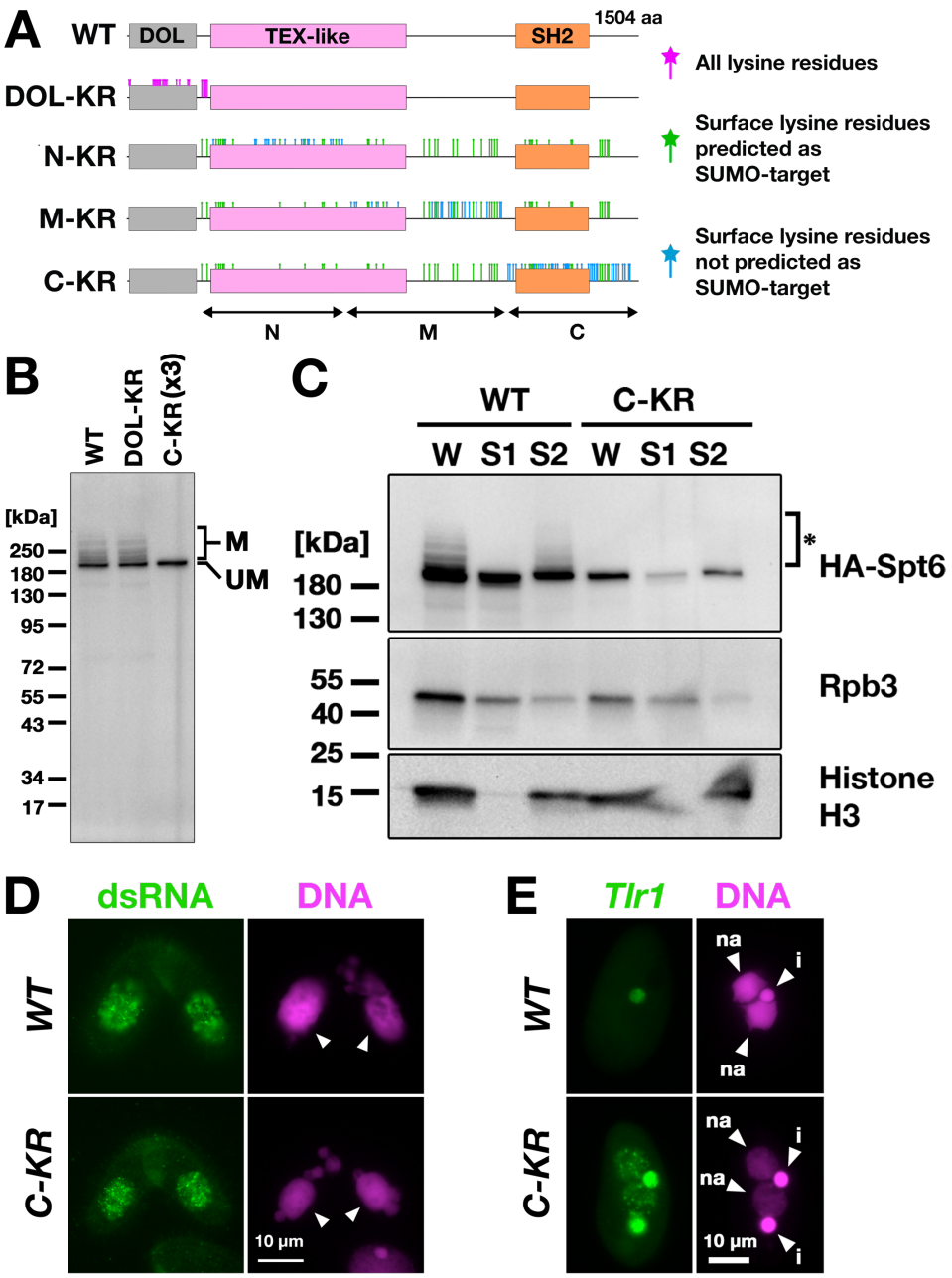


Figure 8. Spt6 SUMOylation is dispensable for parental MAC lncRNA transcription

(A) Schematic representation of *SPT6* mutant constructs with lysine (K)-to-arginine (R) substitutions. Spt6 has a domain of low complexity (DOL) region followed by conserved TEX-like and SH2 domains. The DOL-KR mutant has K to R substitutions for all lysine residues in the DOL region. N-KR, M-KR and C-KR mutants have K to R substitutions for lysine residues that were on the surface of a predicted Spt6 structure in one of the three non-DOL regions (N, M and C-regions) in addition to those surface lysine residues that were also predicted to be SUMOylatable by an algorithm in all three regions.

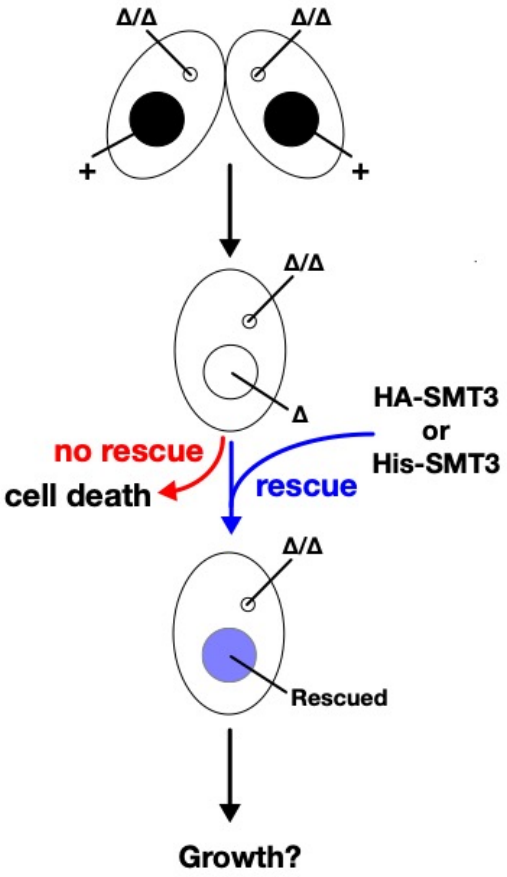
(B) Total proteins were harvested at 4.5 hpm from conjugating *SPT6* KO cells rescued with wild-type (*SPT6-WT*), *SPT6-DOL-KR* or *SPT6-C-KR* constructs. Because mating was low in the SPT6-C-KR rescued cells, 3 times more total protein sample was loaded for these cells. Spt6-HA was detected by western blotting using an anti-HA

antibody. The positions of modified and unmodified Spt6-HA proteins are marked with “M” and “UM”, respectively.

(C) *SPT6-WT* (*WT*) and *SPT6-C-KR* (*C-KR*) rescued cells at 4.5 hpm were analyzed by cell fractionation as described in Figure 7B. The position of SUMOylated Spt6-HA is marked with a bracket.

(D) Accumulation of long dsRNAs in *SPT6-WT* (*WT*) and *SPT6-C-KR* (*C-KR*) rescued cells at 4.5 hpm was analyzed by immunofluorescence staining using the J2 antibody as described in Figure 6B. The parental MACs are marked with arrowheads. All other structures stained by DAPI are MICs. All pictures share the scale bar.

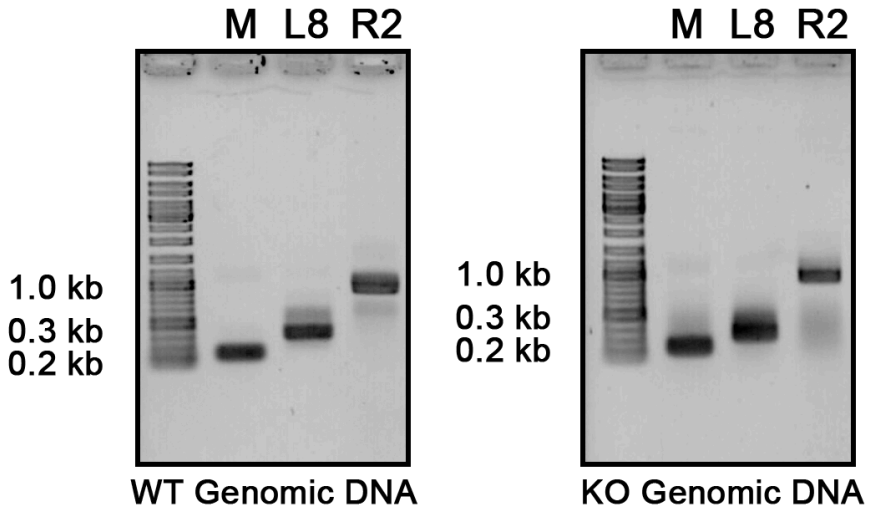
(E) *SPT6-WT* (*WT*) and *SPT6-C-KR* (*C-KR*) rescued cells at 36 hpm were analyzed by DNA-FISH with florescent probes complementary to the Tlr1 element (red). DNA was counterstained with DAPI (blue). The MICs (i) and the new MACs (na) are marked with arrowheads. All pictures share the scale bar.



Constructs	#1	#2
pBCMB1-His	0	0
pBCMB1-His-SMT3	13	-
pBCS3B1-HA	0	0
pBCS3B1-HA-SMT3	16	-

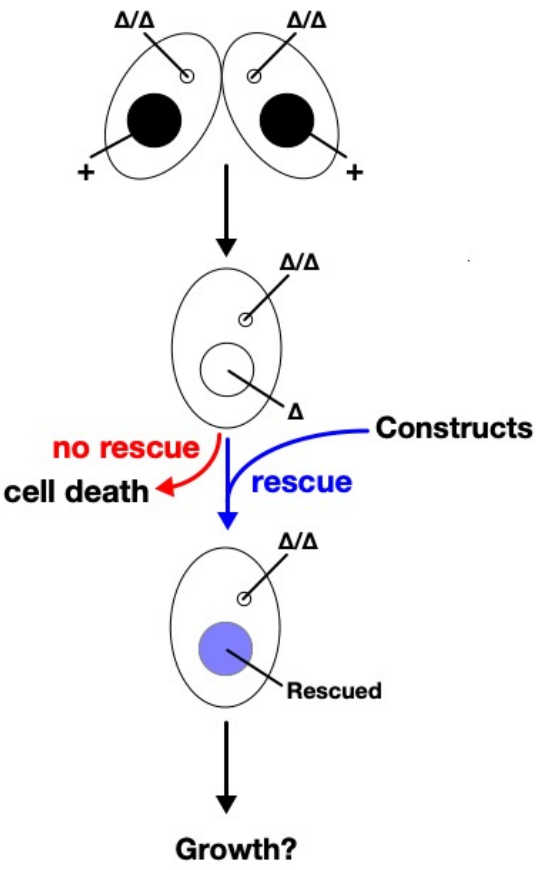
Supplementary Figure 1. HA-Smt3 and His-Smt3 can replace the essential function of Smt3

(Top) Schematic representation of genetic rescue experiment of *SMT3* KO. Two *SMT3* KO homozygous heterokaryon strains were mated and *His-SMT3 or HA-SMT3* constructs (or their corresponding vectors without the SMT3 gene) were introduced into the MAC of their progeny (*SMT3* KO homozygous homokaryon). The progeny survives only when the introduced construct can replace the functions of *SMT3* in vegetative cells. (Bottom) Number of obtained *SMT3* KO homozygous homokaryon strains that were genetically rescued by 10 μg DNA of the indicated constructs are shown. For the empty constructs (pBCMB1-His and pBCS3B1-HA), two transformation experiments (#1 and #2) were done to confirm their inability to restore cell viability of KO cells.



Supplementary Figure 2. Confirmation of the presence of the corresponding genomic loci that were used to detect lncRNAs

Genomic DNA from wild-type (WT) and *EMA2* KO (KO) were used for PCR with the same primers used for the RT-PCR assay in Figure 6A.



Constructs	#1	#2
Empty (pBCSB-HA)	0	0
WT	4	-
DOL-KR	9	-
N-KR	2	2
M-KR	1	4
C-KR	1	3

Supplementary Figure 3. Genetic rescue of *SPT6* KO

(Top) Schematic representation of genetic rescue experiment of *SPT6* KO. Two *SPT6* KO homozygous heterokaryon strains were mated and different *SPT6* expressing constructs (or corresponding pBCSB-HA vector without any SPT6 gene) were introduced into the MAC of their progeny (*SPT6* KO homozygous homokaryon). The progeny survives only when the introduced construct can replace the essential functions of *SPT6* in vegetative cells. (Bottom) Number of obtained *SPT6* KO homozygous homokaryon strains that were genetically rescued by 10 μg DNA of the indicated constructs in two independent attempts (#1 and #2) are shown.