Available online at www.sciencedirect.com

ScienceDirect

European Journal of Protistology 61 (2017) 376–387

European Journal of PROTISTOLOGY

www.elsevier.com/locate/ejop

A comparative study of genome organization and epigenetic mechanisms in model ciliates, with an emphasis on *Tetrahymena*, *Paramecium* and *Oxytricha*

Yurui Wang[a], Yuanyuan Wang[a], Yalan Sheng[a], Jie Huang[a], Xiao Chen[a], Khaled A.S. AL-Rasheid[b], Shan Gao[a,c,*]

[a]*Institute of Evolution & Marine Biodiversity, Ocean University of China, Qingdao 266003, China*
[b]*Zoology Department, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia*
[c]*Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China*

Available online 1 July 2017

Abstract

As a group of unicellular eukaryotes, ciliates offer a unique system to explore epigenetic regulation, mostly due to their nuclear dualism. Ciliates launched a successful radiation after their early evolutionary branching, therefore harboring an unexpectedly rich pool of diverse biological functions and mechanisms. In this review, we compare distinct features of different ciliates in mating type determination, genome organization, DNA methylation, and removal of internal eliminated sequences (IES), with emphasis on *Tetrahymena*, *Paramecium* and *Oxytricha*. Firstly, we review studies on mating type determination in *Paramecium*, one of the foundational phenomena that defined the field of epigenetics, and compare this process with that in *Tetrahymena*. Secondly, we showcase the high diversity in genome structure of several ciliates, such as genome size, gene copy numbers, genome rearrangement, etc. Thirdly, we present a brief description of features and potential functions of 5-methylcytosine (5mC) and N6-methyladenine (6mA) in ciliates so far studied. Fourthly, we describe both the initial and the continuously optimized scan RNA (scnRNA) model for IES elimination in *Tetrahymena* and contrast it with that in *Paramecium* and *Oxytricha*. Finally, we discuss the importance of integrative approaches to the study of epigenetic diversity in ciliates and provide possible directions for future research.
© 2017 Elsevier GmbH. All rights reserved.

Keywords: Ciliates; Epigenetics; Integrative biology; *Tetrahymena thermophila*

Introduction

Epigenetic modifications, which include heritable changes that are not based on primary DNA sequences, are mainly of three important kinds: RNA interference, histone post-translational modifications, and DNA methylation (Holliday and Pugh, 1975; Russo et al., 1996; Volpe et al., 2002). These modifications work together with established genetic information to shape the cellular environment, thus governing almost every aspect of biological processes.

Ciliates are a large and diverse group of single-celled eukaryotes distributed throughout the world (Foissner, 1995; Kahl, 1931; Lynn, 2008; Lynn and Small, 2002; Song et al., 2009). Each ciliate has two functionally distinct nuclei (Hammerschmidt et al., 1996; Nanney and Caughey, 1953). The ampliploid macronucleus (MAC) stores the somatic

*Corresponding author at: Institute of Evolution & Marine Biodiversity, Ocean University of China, Qingdao 266003, China.
E-mail address: shangao@ouc.edu.cn (S. Gao).

http://dx.doi.org/10.1016/j.ejop.2017.06.006
0932-4739/© 2017 Elsevier GmbH. All rights reserved.

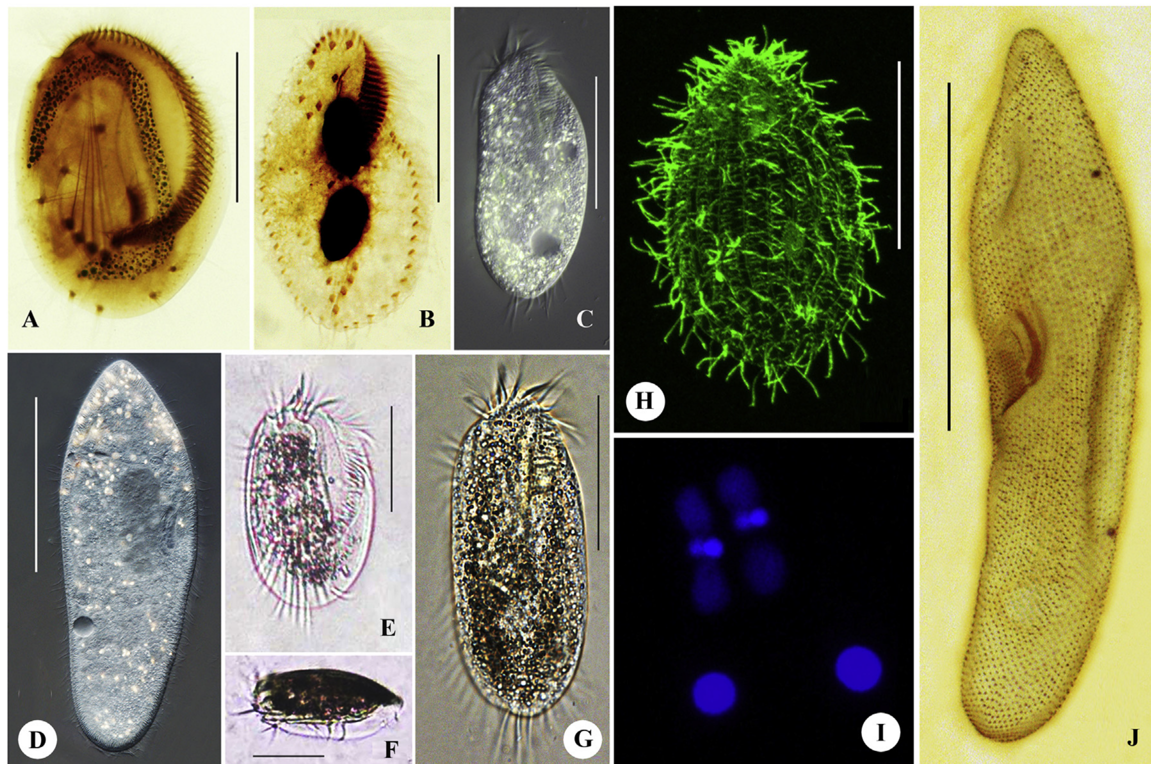

Fig. 1. Photomicrographs of four model ciliates: two spirotrichs—the euplotian *Euplotes harpa* (A, E, F), hypotrich *Oxytricha trifallax* (B, C, G), and two oligohymenophoreans—*Paramecium caudatum* (D, J) and *Tetrahymena thermophila* (H, I). Scale Bars: 60 μm (A, E, F), 50 μm (B, C, G), 20 μm (H), 100 μm (D, J).

genome and governs the cellular phenotype, while the presumably diploid micronucleus (MIC) contains a germline genome and is transcriptionally silent during vegetative growth. Ciliates have served as model organisms in many research fields, including cytology, zoology, evolutionary biology and genetics (Hausmann and Bradbury, 1996). In addition, ciliates provide novel insight into epigenetic regulation due to their nuclear dimorphism (Hammerschmidt et al., 1996; Nanney and Caughey, 1953). Ciliates display a high level of diversity not only in morphology (Fig. 1), but also in genetic and epigenetic regulations. Here, we compare distinct features in different ciliates, with emphasis on *Tetrahymena thermophila*, *Paramecium "aurelia"* and *Oxytricha* sp.

Mating type determination

Besides asexual reproduction, ciliates also have a sexual stage, conjugation, to confer physiological regeneration and increase genotypic variability (Rogers and Karrer, 1985). During conjugation, a newly-formed zygotic MIC divides and then differentiates into progeny MACs and MICs while accompanying parental MACs disintegrate.

Mating types were first discovered in the *Paramecium "aurelia"* complex by Sonneborn (1937). Two complementary mating types—O (odd) and E (even)—exist in the *Paramecium "aurelia"* complex. Notably, the O/E alternative of *P. tetraurelia* is maternally inherited, a signal discovery in epigenetics (Hiwatashi, 1968). Three genes, *mtA, mtB* and *mtC*, are required for type E expression (Singh et al., 2014). Type E is determined by the expression of the transmembrane protein mtA, which plays a key role in species-specific recognition (Singh et al., 2014). Meanwhile, *mtB* and *mtC* gene products seem to be required for the *mtA* expression (Singh et al., 2014). The *mtA* promoter is found to be excised in mating type O. This excision is mediated by small RNAs in a similar way to the deletion of Internal Eliminated Sequences (IESs), which will be introduced in a following section (Singh et al., 2014). A similar mating type determination model is also found in *P. octaurelia* (Singh et al., 2014). In contrast, a different mating type switch is proposed in *P. septaurelia*: its O expression is determined by the excision of *mtB* coding-sequence segments, which also causes *mtA* gene silencing (Singh et al., 2014).

Tetrahymena thermophila has seven mating types (Nanney and Caughey, 1953). After conjugation, the progeny cells can possess any one of the seven mating types. In other words, its mating type is determined randomly (Doerder et al., 1995). Two adjacent genes, *MTA* and *MTB*, are arranged head-to-head to constitute the somatic mating type locus (Cervantes et al., 2013). Each gene can encode a mating type-specific segment and a transmembrane (TM) domain, which is highly conserved in all mating types (Cervantes et al., 2013). The germline mating type locus contains a tan-

dem array of incomplete gene pairs, one representing each mating type (Cervantes et al., 2013). During the progeny somatic differentiation, one gene pair is chosen and completely assembled while the remaining gene pairs are deleted (Cervantes et al., 2013). Cervantes et al. (2013) also proposed that programmed DNA rearrangements are involved in its mating type determination process.

Genome structure

Within the past decade, micronuclear/macronuclear genome sequences of several ciliates have been reported (Arnaiz et al., 2012; Aury et al., 2006; Chen et al., 2014; Eisen et al., 2006; Hamilton et al., 2016; Slabodnick et al., 2017; Swart et al., 2013). Genome studies have made a major impact on ciliate research and confirmed the high diversity among ciliates that was previously expected by single gene sequences (Table 1).

T. thermophila possesses one macronucleus (MAC) and one micronucleus (MIC), the genome structures of which are dramatically different (Chen et al., 2016; Zhao et al., 2016) (Table 2). The MIC genome is approximately 157 Mb while the MAC genome is about 103 Mb (Eisen et al., 2006; Hamilton et al., 2016). In terms of chromosome numbers, the MIC contains only five pairs of chromosomes and divides mitotically in vegetative cells, while the MAC consists of nearly 181 centromere-less chromosomes and divides amitotically without spindle formation and chromosome condensation (Hamilton et al., 2016; Ray, 1956; Xiong et al., 2012). Notably, an 11-kb single-copy micronuclear rDNA is duplicated as a 21-kb palindromic minichromosome and amplified to ∼10,000 copies in the MAC (Yao and Gall, 1977; Yao and Gorovsky, 1974). Meanwhile, some histone modifications and variants, like histone acetylation, H3K4 methylation, H3.3 and H2A.Z, are exclusively distributed in the MAC and signal different transcriptional activities (Allis et al., 1980; Cervantes et al., 2006; Cui et al., 2006; Henikoff and Smith, 2015; Strahl et al., 1999).

The *Tetrahymena* genome undergoes programmed chromosome breakage and extensive DNA elimination during development of the new MAC (Orias et al., 2011). In this process, the five MIC chromosomes are fragmented into hundreds of somatic minichromosomes at a 15-bp chromosome breakage sequence (Cbs), which is only found in tetrahymenine ciliates (Blackburn and Gall, 1978; Hamilton et al., 2016; Lin et al., 2016; Yao et al., 1990). Chromosome breakage sequences share a highly conserved motif—WAAACCAACCYCNHW (W = A/T; Y = T/C; H = A/T/C; N = G/A/T/C)—in *T. thermophila* and other *Tetrahymena* spp. (Hamilton et al., 2016). Some minichromosomes are eliminated about 10 doubling times after conjugation while the rest are maintained in the mature MAC genome (Lin et al., 2016), providing a distinct method of genome reduction (Lin et al., 2016). Internal Eliminated Sequences (IESs) comprise approximately 30% of the MIC genome (Coyne et al., 2012). These IESs are deleted by a small RNA-induced heterochromatin formation (Liu et al., 2007; Mochizuki et al., 2002; Taverna et al., 2002), while the remaining Macronucleus Destined Sequences (MDSs) are ligated to form the MAC minichromosomes. Subsequently, endoreplication follows in the differentiating MAC to increase its copy number from 2 to ∼45 (Doerder et al., 1992).

Paramecium tetraurelia belongs to the same class Oligohymenophorea as *Tetrahymena* (Adl et al., 2012; Lynn, 2008). Nevertheless, their respective genome structures differ markedly as follows: (1) the genome of *P. tetraurelia* has undergone three whole-genome duplication (WGD) events, with the most recent one occurring after the divergence of the *P. "aurelia"* complex from *P. caudatum* and *P. multimicronucleatum* (Aury et al., 2006), while WGD has not been detected in the *Tetrahymena* genome; (2) the copy number of their MAC chromosomes is different (*P. tetraurelia* is 800C while *T. thermophila* is 45C) (Aury et al., 2006; Doerder et al., 1992); (3) some short IESs of *P. tetraurelia* lie in protein-coding exons and require precise elimination to guarantee proper translation (Arnaiz et al., 2012), while the vast majority of IESs of *Tetrahymena* lie in intergenic or intronic regions with variable junction sites and do not require precise elimination (Bétermier, 2004; Fass et al., 2011; Hamilton et al., 2016).

Oxytricha belongs to another ciliate class—the Spirotrichea (Adl et al., 2012; Lynn, 2008). Its genetic distance from *Tetrahymena* is comparable to that between humans and fungi (Parfrey et al., 2011), and this distance is demonstrated by significant differences in genome features. MAC chromosomes of *Oxytricha* are highly fragmented, a feature shared by other spirotrichs (Jahn and Klobutcher, 2002). Their average size is 3.2 kb and ∼90% are "nanochromosomes" that encode a single gene (Swart et al., 2013). The copy number of these nanochromosomes is quite variable, averaging ∼2000C, which is considerably higher than that of *Tetrahymena* and *Paramecium* (Prescott, 1994).

Table 1. Differences between epigenetic mechanisms in ciliate model species.

	Tetrahymena thermophila	Paramecium tetraurelia	Oxytricha trifallax	Euplotes crassus
MAC genome size	103 Mb[1]	72 Mb[15]	50 Mb[25]	∼50 Mb?[26]
MIC genome size	157 Mb[2]	100 Mb[16]	∼500 Mb[38]	–
Whole genome duplication	No	Yes[15]	No	No
MAC chromosome numbers	181[3]	∼200[15]	15,600[26]	–
MIC chromosome numbers	5[4]	≥50[15]	–	–

Table 1 (*Continued*)

	Tetrahymena thermophila	Paramecium tetraurelia	Oxytricha trifallax	Euplotes crassus
MAC ploidy			Variable	Variable
	45[5]	800[15]	~2,000[27]	~2000[34]
MIC ploidy	2[4]	2[17]	2[17]	2[17]
Nano-chromosome	No	No	Yes[26]	Yes[35]
Genome unscrambling	No	No	Yes[28]	No
IES percentage	~30%[6]	~30%[16]	90%[38]	95%[18]
IES location (genic/intergenic)	Intergenetic[7]	Both, mainly genetic[16,18]	Both, mainly genetic[29]	–
IES classification	Maternally regulated[8,a]	Both[19,20]	–	–
IES deletion mechanisms (precise/imprecise)	Imprecise[9]	Precise[21]	Precise[30]	–
scnRNA source (IES/MDS)	IES and MDS[10]	IES and MDS[22]	MDS[25]	–
Late scnRNA production (source)	New macronucleus[11]	New macronucleus[37]	–	–
Pointer	No	No	Yes[31]	No
Mating type	7[12]	2[23]	9[32]	38[36]
DNA methylation	6mA[13]	5mC?[24,b]	5mC[33]	–
Number of genes	24,700[2]	39,642[15]	~18,400[26]	–
Number of IES	6000[14]	65,000[19]	150,000[30]	–

[1] Eisen et al. (2006).
[2] Hamilton et al., (2016).
[3] Xiong et al., (2012).
[4] Ray, (1956).
[5] Doerder et al., (1992).
[6] Coyne et al., (2012).
[7] Schoeberl et al., (2012).
[8] Chalker and Yao, (1996).
[9] Fass et al., (2011).
[10] Chalker and Yao, (2001).
[11] Noto et al., (2015).
[12] Doerder et al., (1995).
[13] Gorovsky et al., (1973).
[14] Coyne et al., (1996).
[15] Aury et al., (2006).
[16] Arnaiz et al., (2012).
[17] Ammermann, (1971).
[18] Jahn and Klobutcher, (2002).
[19] Duharcourt et al., (1995).
[20] Duharcourt et al., (1998).
[21] Lhuillier-Akakpo et al., (2014).
[22] Duharcourt et al., (2009).
[23] Sonneborn, (1937).
[24] Kwok and Ng, (1989).
[25] Fang et al., (2012).
[26] Swart et al., (2013).
[27] Prescott, (1994).
[28] Greslin et al., (1989).
[29] Prescott et al., (2002).
[30] Nowacki et al., (2008).
[31] Allis et al., (1987).
[32] Siegel, (1956).
[33] Bracht et al., (2012).
[34] Bender and Klein, (1997).
[35] Klobutcher, (1999).
[36] Dini and Luporini, (1979).
[37] Sandoval et al., (2014).
[38] Chen et al., (2014).
[a] Almost all *Tetrahymena* IESs tested so far are maternally regulated.
[b] 5mC is not biochemically detected. Mac and Mic represent macronucleus and micronucleus, respectively; IES, internal eliminated sequences; MDS, macronucleus destined sequences.

Table 2. Differences between the macronucleus (MAC) and micronucleus (MIC) in *Tetrahymena thermophila.*

	MAC	MIC	Literature
Ploidy	45	2	(Doerder et al., 1992; Ray, 1956)
Chromosome numbers	181	5	(Ray, 1956; Xiong et al., 2012)
Genome size (Mb)	103	157	(Eisen et al., 2006; Hamilton et al., 2016)
Mitosis[a]	No	Yes	(Gorovsky et al., 1978)
Linker histone	Hho lp	micLH	(Allis et al., 1984; Hayashi et al., 1987; Wu et al., 1986)
H2A.Z	Yes	No	(Henikoff and Smith, 2015)
H3.3	Yes	No	(Henikoff and Smith, 2015)
Cnp1p	No	Yes	(Cervantes et al., 2006; Cui and Gorovsky, 2006)
Histone acetylation	Yes	No	(Cervantes et al., 2006; Cui and Gorovsky, 2006)
H3K4 methylation	Yes	No	(Strahl et al., 1999)
H3S10 phosphorylation	No	Yes	(Allis and Gorovsky, 1981)
Transcription	Yes	No[b]	(Gorovsky et al., 1978)

[a]Mitosis during vegetative growth.
[b]No transcription during vegetative growth. Y and N represent Yes and No, respectively.

Furthermore, the MIC genome is highly complex, containing scrambled genes that require elaborate chromosome rearrangements (e.g., deletion, inversion and permutation) during MAC development in order to generate a functional MAC genome (Chen et al., 2014). The MIC genome of *Stylonychia*, another oxytrichid spirotrich like *Oxytricha*, also has scrambled genes (Prescott, 1999). In *Oxytricha*, IESs make up nearly 90% of the MIC genome, and, as in the oligohymenophoreans, are eliminated during MAC differentiation. This produces a small-sized MAC genome (50 Mb) from a medium-sized MIC genome (∼500 Mb) (Chen et al., 2014; Nowacki et al., 2008; Swart et al., 2013). The long non-coding RNAs (lncRNAs), derived from the parental MAC genome, have been shown to act as templates to guide DNA unscrambling (Nowacki et al., 2008).

In another spirotrich, *Euplotes*, the MAC genome is also highly fragmented, and nanochromosomes make up a large portion (Klobutcher, 1999). *Euplotes* has two translational peculiarities that can distinguish it from other ciliates: stop codon usage and +1 programmed ribosomal frameshift (PRF). The stop codon UGA is used in *Tetrahymena* and *Paramecium* while in *Euplotes* it is reassigned as the cysteine or selenocysteine codon. The stop codons UAA and UAG in *Tetrahymena* and *Paramecium* encode glutamine while *Euplotes* uses both as stop codons (Salim et al., 2008). *Euplotes* is so far the only ciliate reported to require +1 PRF: a number of genes of euplotids have been reported to require this frameshift to express a functional protein (e.g., the *E. aediculatus* La motif protein (p43) and the *E. crassus* Tec2 transposon ORF2 protein), while 11% of genes in *E. octocarinatus* require the +1 PRF (Aigner et al., 2000; Doak et al., 2003; Jahn et al., 1993; Wang et al., 2016). Compared to other model ciliates, molecular genetic tools in the *Euplotes* system are far less powerful. Gene silencing in *Euplotes* spp. can be achieved by feeding them *E. coli* expressing dsRNA corresponding to a target locus (Galvani and Sperling, 2002), but a transformation system is urgently needed to enable more systematic studies on *Euplotes* spp.

Distantly related to the four model ciliate genera discussed above, the heterotrich *Stentor coeruleus* is a giant cell (up to 4000 µm long when extended) that possesses a moniliform MAC with a variable number of nodules in each cell (Foissner et al., 1992). This huge cell size is correlated with an extremely high copy number of its MAC genes (average 60,000C) (Slabodnick et al., 2017). The genome of *Stentor* is distinct in at least two aspects. First, it contains extremely short introns (15–16 nt) with a conserved motif $GTAN_5TAN_3AG$ (N = G/A/T/C) (Slabodnick et al., 2017). These tiny introns, also found in other heterotrichs, are shorter than any other introns yet reported (Slabodnick et al., 2017). More interestingly, *Stentor* uses a standard genetic code instead of a "ciliate code" (Swart et al., 2016), indicating that the deviation of genetic codes occurred after the heterotrichs, like *Stentor*, diverged from other ciliates (Slabodnick et al., 2017). In another heterotrich *Condylostoma magnum* and a karyorelictid *Parduczia* sp., all three stop codons could be reassigned as a sense codon (Swart et al., 2016). Translation termination relies on the mRNA 3′ ends in a context-dependent manner (Swart et al., 2016).

DNA methylation

DNA methylation, one of the most important epigenetic modifications, can modulate gene expression by influencing the interaction between DNA and transcription factors (Fig. 2) (Feng et al., 2010; Smith and Meissner, 2013). Common DNA methylations include 5-methylcytosine (5mC), N6-methyladenine (6mA), and N4-methylcytosine (4mC) (Cheng, 1995; Ratel et al., 2006; Wion and Casadesús, 2006). 5mC has been most intensively investigated as it plays a significant role in epigenetics in other eukaryotes. In contrast, 6mA was only detected in prokaryotes and some lower eukaryotes (Hattman et al., 1978; Kay et al., 1994; Reyes et al., 1997) until some recent studies characterized its presence and potential functions in multicellular eukaryotes (Fu

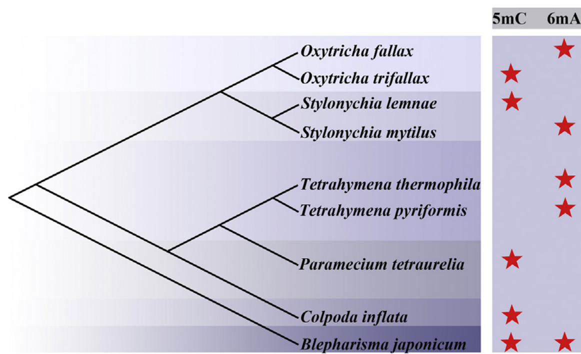

Fig. 2. Phyletic patterns of methylation of DNA in ciliates as 5-methylcytosine (5mC) and 6-methyladenosine (6 mA). Ciliate species are shown according to their positions in a consensus phylogram based on (Gao et al., 2016). Red stars represent the presence of DNA methylation in species.

et al., 2015; Greer et al., 2015; Koziol et al., 2016; Liu et al., 2016; Wu et al., 2016; Zhang et al., 2015). Eukaryotic 5mC and 6mA are distinctly different. For instance, 5mC is mainly located in nucleosomal regions (Razin and Cedar, 1977) while 6mA is enriched in linker DNA (Harrison et al., 1986). Both 5mC and 6mA are associated with transcription inactivation (Tajima and Suetake, 1998; Wu et al., 2016), while 6mA can also be associated with transcription activation (Fu et al., 2015; Greer et al., 2015 Greer et al., 2015). 5mC and 6mA are apparently sporadically distributed in different clades of ciliates (Fig. 2).

6mA is the only form of DNA methylation in *T. thermophila* (Blackburn et al., 1983; Hattman et al., 1978). In vegetative cells, approximately 0.8% of the MAC adenine is methylated, while 6mA is undetectable in the MIC (Blackburn et al., 1983; Hattman et al., 1978). The level of 6mA is relatively stable in different stages of vegetative growth or when compared with starved cells (Gorovsky et al., 1973; Karrer and VanNuland, 2002). During conjugation, *de novo* 6mA can be detected during new MAC development, at about 14–16 h after the initiation of conjugation (Harrison et al., 1986). It has also been reported that 6mA preferentially locates in linker DNA regions in the sequence of NAT (N = G/A/T/C) (Hattman et al., 1978). In its congener *Tetrahymena pyriformis*, 6mA is also the only detected methylated base (Hattman et al., 1978).

6mA content varied slightly in the *Paramecium "aurelia"* complex, ranging from 2.1% to 2.5% (Cummings et al., 1974). It is proposed that 6mA may be involved in the transcription inactivation of some sequences in MAC, but this requires further validation (Kwok and Ng, 1989). 5mC has been indirectly discovered in *P. tetraurelia* using the nucleoside analogue 5-azacytidine and is proposed to modulate the programming of the developing MAC (Kwok and Ng, 1989).

6mA was the only detected methylation in *O. fallax* (Rae and Spear, 1978), yet a more recent study in *Oxytricha trifallax* detected 5mC directly by mass spectrometry (Bracht et al., 2012). As proposed by these latter authors, 5mC occurs only during nuclear development and genome rearrangement where it marks sequences within transposon elements, including parental MAC chromosomes, MIC-limited repetitive elements, and aberrantly-spliced gene rearrangement products, indicating that 5mC might play key roles in DNA elimination during genome rearrangements (Bracht et al., 2012).

As in *O. trifallax*, 5mC was detected within transposable elements during MAC differentiation of *Stylonychia lemnae* (Juranek et al., 2003). 5mC is also present at low levels in the vegetative MIC in the sequence of CCWGG (W = A/T). Juranek et al. (2013) accordingly proposed that 5mC may be involved in silencing sequences and in heterochromatin formation instead of in MAC differentiation. In contrast, 6mA is the only detectable methylated base in *Stylonychia mytilus* (Ammermann et al., 1981).

Only 5mC is detected in the colpodid *Colpoda inflata* and this methylation pattern changes during the encystment process, indicating that 5mC participates in the regulation of gene expression during encystment (Palacios et al., 1994; Taylor et al., 1983). In the heterotrich *Blepharisma japonicum*, both 5mC and 6mA were observed in the MAC of vegetative cells (Salvini et al., 1986; Salvini et al., 1984). 6mA was proposed to be play a role in gene activation in *Blepharisma* (Salvini et al., 1986).

IES elimination mediated by small RNA

In ciliated protozoa, small RNAs play key roles in IES excision in germline and somatic genomes. This pathway has been well studied in three ciliate genera—*Paramecium, Tetrahymena* and *Oxytricha* (Fang et al., 2012; Mochizuki et al., 2002; Noto et al., 2015; Sandoval et al., 2014) (Fig. 3).

The original scan RNA (scnRNA) model in *Tetrahymena* can be summarized as follows: (1) bidirectional transcripts are synthesized by RNA polymerase II during parental MIC meiosis (Chalker and Yao, 2001); (2) Dicer-like ribonuclease (encoded by *DCL1*) cuts long transcripts to short double-strand RNA (~27–30 bp), the so called scnRNAs (Malone et al., 2005; Mochizuki and Gorovsky, 2005); (3) scnRNAs are exported to the cytoplasm and form a complex by associating with the Argonaute (Ago) family protein (encoded by *TWI1*) (Mochizuki et al., 2002); (4) Ago-scnRNA complexes are then transferred into the parental MAC and compared with nascent non-coding transcripts from the MAC genome. scnRNAs that can pair with MAC transcripts (MDS) are destroyed, and the refined pool of scnRNAs is now complementary to the IESs of the MIC (Mochizuki et al., 2002); (5) the remaining Ago-scnRNA complexes are exported to the developing MAC of the progeny and initiate the second round of the scanning process (Mochizuki et al., 2002), targeting IES regions for heterochromatin formation and subsequent elimination (Liu et al., 2007; Taverna et al., 2002).

This model has been continually improved, largely owing to the works of the Mochizuki laboratory. First, it was con-

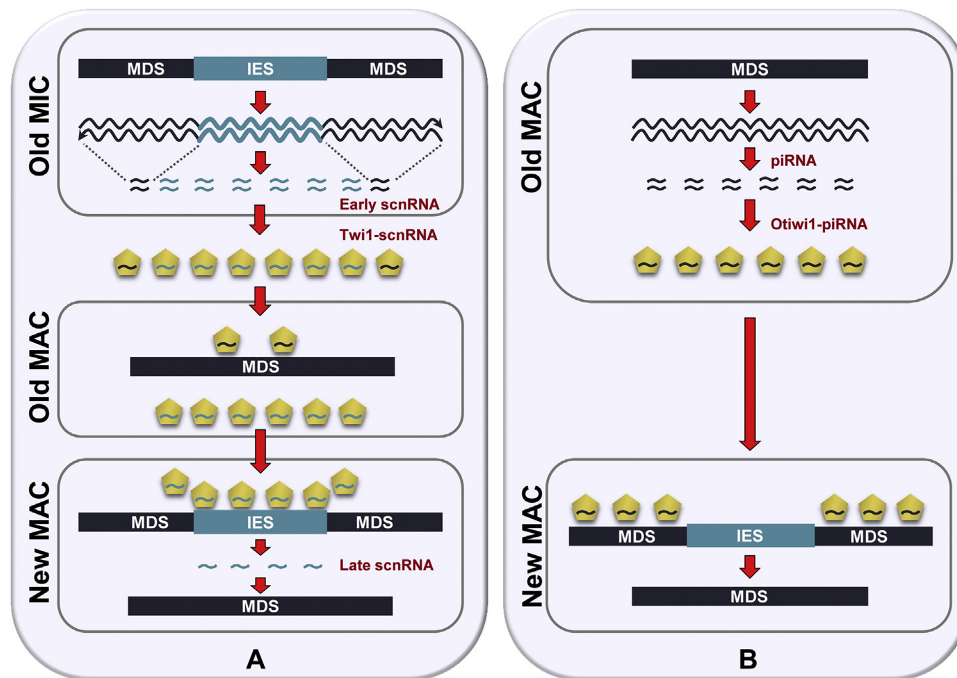

Fig. 3. Small RNAs play different roles during differentiation of a new macronucleus (MAC) in *Tetrahymena, Paramecium*, and *Oxytricha.* (A) The scan RNA (scnRNA) model of *Tetrahymena* and *Paramecium* in which scnRNAs are generated from the MIC genome and cause degradation of internally eliminated sequences (IESs) during development of the new MAC. (B) The Piwi-interacting small RNAs (piRNA) model in *Oxytricha* in which piRNAs are generated from the parental MAC genome and protect macronucleus-destined sequences (MDS) from degradation. Modified from (Fang et al., 2012; Orias et al., 2011).

sidered that scnRNAs are produced in an unbiased way in the MIC genome, including both the IES and MDS regions (Mochizuki et al., 2002). Subsequently however, Schoeberl et al. (2012) found that scnRNAs are preferentially produced from IESs in the parental MIC. Second, scnRNAs were thought to be produced exclusively from the parental MIC during early conjugation (Mochizuki et al., 2002), whereas Noto et al. (2015) demonstrated that in addition to "Early-scnRNAs" in the MIC, "Late-scnRNAs" are produced in the progeny MAC and derived from IESs prior to their excision during late conjugation. Third, instead of only recognizing IESs from which they are derived (Mochizuki et al., 2002), Noto et al. (2015) showed that scnRNAs can also recognize other IESs *in trans* and considered this as a genome-wide *trans*-recognition network, which can guarantee efficient elimination of IESs in the genome of *Tetrahymena*. Nevertheless, a great deal of work is still needed before we fully understand what mechanisms govern the preferential transcription in the parental MIC and how the boundaries are determined for deletion of the IESs in the progeny MAC.

A similar scnRNA model has been proposed in *Paramecium* (Bouhouche et al., 2011; Lepère et al., 2009; Sandoval et al., 2014). However, there are several differences between *Paramecium* and *Tetrahymena*: (1) *Paramecium* contains maternally and non-maternally controlled IESs, while almost all the IESs studied so far in *Tetrahymena* are maternally controlled (Chalker and Yao, 1996; Duharcourt et al., 1995; Duharcourt et al., 1998); (2) *Paramecium* scnRNAs and IES-derived small interfering RNAs (iesRNAs) are produced by different Dicer-like proteins (Noto et al., 2015; Sandoval et al., 2014) while in *Tetrahymena* Early- and Late-scnRNAs are processed by the same Dicer-like ribonuclease—DCL1; (3) *Paramecium* iesRNAs are transcribed and processed from excised IESs while *Tetrahymena* Late-scnRNAs are produced before the IESs are deleted from developing MAC chromosomes (Noto et al., 2015; Sandoval et al., 2014); (4) most *Paramecium* IESs can express both scnRNAs and iesRNAs while some *Tetrahymena* IESs can only express Late-scnRNAs (Noto et al., 2015; Sandoval et al., 2014); (5) ~98% IESs of *Paramecium* are unique (Arnaiz et al., 2012), and hence, instead of being involved in *trans*-recognition between different IESs as in *Tetrahymena*, iesRNAs of *Paramecium* can recognize other copies of the same IESs to ensure IESs elimination (Noto et al., 2015; Sandoval et al., 2014).

Oxytricha employs an entirely different way to eliminate IESs: (1) scnRNAs in *Tetrahymena* and *Paramecium* are generated from the parental MIC genome during early conjugation, whereas Piwi-interacting small RNAs (piRNAs) in *Oxytricha* derive from the parental MAC (Fang et al., 2012); (2) the refined pool of scnRNAs in *Tetrahymena* and *Paramecium* target IESs for deletion while piRNAs in *Oxytricha* protect the MDS for retention (Fang et al., 2012; Mochizuki et al., 2002). This indicates that scnRNAs behave in a cost-efficient way by targeting a minority of the genome: ~10% of the genome is MDS in *Oxytricha* compared to ~30% of

the genome as IESs in *Tetrahymena* and *Paramecium* (Arnaiz et al., 2012; Chen et al., 2014; Coyne et al., 2012; Fang et al., 2012; Prescott, 1994); (3) piggyBac transposase-related proteins are required in IES elimination of both *Tetrahymena* and *Paramecium* while germline-limited transposons mediate their own deletion in *Oxytricha* (Baudry et al., 2009; Cheng et al., 2010; Nowacki et al., 2009; Shieh and Chalker, 2013).

Conclusion and perspectives

Studies of ciliates have contributed deep insights into epigenetic regulation. The discovery of scnRNA and histone methylation-mediated DNA elimination not only illuminates the mystery of MAC-MIC differentiation in ciliates, but also provides mechanisms for epigenetic regulation in higher eukaryotes. However, some questions remain unanswered and merit future study. For example, what mechanism underlies the replication of histone modifications to achieve the transgenerational inheritance of chromatin states; how do scnRNAs preferentially transcribe from IES regions in the meiotic MIC and what defines the deletion boundaries of IESs in the new MAC to maintain a stable genome structure at the population level.

Ciliates are an extremely diverse group, not only morphologically, but also in genome structure and in their mechanisms of epigenetic regulation. More extensive investigations should be carried out on non-model ciliates to explore this rich pool of diversity. To accomplish this task, integrative approaches should be applied for the following reasons. First, to study either epigenetic diversity of different species or phenotypic plasticity of one species, accurate taxonomic identification and ecological role characterization are required to guarantee proper identification from the outset. For example, the proper identification of three closely related and morphologically similar species, *Oxytricha trifallax*, *Oxytricha fallax*, and *Sterkiella histriomuscorum*, has been crucial to ensure the correct interpretation of results (Foissner and Berger, 1999; Zoller et al., 2012). Second, genetic toolkits are required to pinpoint key players in pathways and to reveal their functions and the underlying mechanisms. Last but not least, phylogenetic studies, both to determine the systematic relationships of taxa but also to explore the evolutionary conservation/flexibility of protein families, will provide a deeper understanding of adaptive evolution in the diverse lineages of ciliates.

Acknowledgements

We thank Ms. Chunyu Lian and Mr. Mingjian Liu, graduate students in the Laboratory of Protozoology, Ocean University of China (OUC), for their help in species identification. We thank the reviewers for insightful critiques and Dr. William A. Bourland, Boise State University, USA, for English editing. Our thanks are also due to Prof. W. Song, OUC, the China-based leading scientist of the IRCN-BC program, for his kind suggestions during the drafting of this manuscript. This work was supported by the Natural Science Foundation of China (31522051), AoShan Talents Program by Qingdao National Laboratory for Marine Science and Technology (2015ASTP), the International Research Coordination Network for Biodiversity of Ciliates (IRCN-BC) (31111120437), a postdoctoral grant from Shandong Province (201501014) and Deanship of Scientific Research, King Saud University, Prolific research group (PRG-1436-01).

References

Adl, S.M., Simpson, A.G., Lane, C.E., Lukeš, J., Bass, D., Bowser, S.S., Brown, M.W., Burki, F., Dunthorn, M., Hampl, V., Heiss, A., Hoppenrath, M., Lara, E., Le, G.L., Lynn, D.H., McManus, H., Mitchell, E.A.D., Mozley-Stanridge, S.E., Parfrey, L.W., Pawlowski, J., Rueckert, S., Shadwick, L., Schoch, C., Smirnov, A., Spiegelt, F.W., 2012. The revised classification of eukaryotes. J. Eukaryot. Microbiol. 59, 429–514.

Aigner, S., Lingner, J., Goodrich, K.J., Grosshans, C.A., Shevchenko, A., Mann, M., Cech, T.R., 2000. *Euplotes* telomerase contains an La motif protein produced by apparent translational frameshifting. EMBO J. 19, 6230–6239.

Allis, C.D., Allen, R.L., Wiggins, J.C., Chicoine, L.G., Richman, R., 1984. Proteolytic processing of H1-like histones in chromatin: a physiologically and developmentally regulated event in *Tetrahymena* micronuclei. J. Cell Biol. 99, 1669–1677.

Allis, C.D., Colavito-Shepanski, M., Gorovsky, M.A., 1987. Scheduled and unscheduled DNA synthesis during development in conjugating *Tetrahymena*. Dev. Biol. 124, 469–480.

Allis, C.D., Glover, C.V., Bowen, J.K., Gorovsky, M.A., 1980. Histone variants specific to the transcriptionally active, amitotically dividing macronucleus of the unicellular eucaryote, *Tetrahymena thermophila*. Cell 20, 609–617.

Allis, C.D., Gorovsky, M.A., 1981. Histone phosphorylation in macro- and micronuclei of *Tetrahymena thermophila*. Biochemistry 20, 3828–3833.

Ammermann, D., 1971. Morphology and development of the macronuclei of the ciliates *Stylonychia mytilus* and *Euplotes aediculatus*. Chromosoma 33, 209–238.

Ammermann, D., Steinbrück, G., Baur, R., Wohlert, H., 1981. Methylated bases in the DNA of the ciliate *Stylonychia mytilus*. Eur. J. Cell Biol. 24, 154.

Arnaiz, O., Mathy, N., Baudry, C., Malinsky, S., Aury, J.-M., Wilkes, C.D., Garnier, O., Labadie, K., Lauderdale, B.E., Le Mouël, A., 2012. The *Paramecium* germline genome provides a niche for intragenic parasitic DNA: evolutionary dynamics of internal eliminated sequences. PLoS Genet. 8, e1002984.

Aury, J.-M., Jaillon, O., Duret, L., Noel, B., Jubin, C., Porcel, B.M., Ségurens, B., Daubin, V., Anthouard, V., Aiach, N., 2006. Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. Nature 444, 171–178.

Bétermier, M., 2004. Large-scale genome remodelling by the developmentally programmed elimination of germ line sequences in the ciliate *Paramecium*. Res. Microbiol. 155, 399–408.

Baudry, C., Malinsky, S., Restituito, M., Kapusta, A., Rosa, S., Meyer, E., Bétermier, M., 2009. PiggyMac, a domesticated

piggyBac transposase involved in programmed genome rearrangements in the ciliate *Paramecium tetraurelia*. Genes Dev. 23, 2478–2483.

Bender, J., Klein, A., 1997. The telomere binding protein of *Euplotes crassus* prevents non-specific transcription initiation but has no role in positioning transcription initiation complexes. Nucleic Acids Res. 25, 2877–2882.

Blackburn, E.H., Gall, J.G., 1978. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*. J. Mol. Biol. 120, 33–53.

Blackburn, E.H., Pan, W.-C., Johnson, C.C., 1983. Methylation of ribosomal RNA genes in the macronucleus of *Tetrahymena thermophila*. Nucleic Acids Res. 11, 5131–5145.

Bouhouche, K., Gout, J.F., Kapusta, A., Bétermier, M., Meyer, E., 2011. Functional specialization of Piwi proteins in *Paramecium tetraurelia* from post-transcriptional gene silencing to genome remodelling. Nucleic Acids Res. 39, 4249–4264.

Bracht, J.R., Perlman, D.H., Landweber, L.F., 2012. Cytosine methylation and hydroxymethylation mark DNA for elimination in *Oxytricha trifallax*. Genome Biol. 13, R99.

Cervantes, M.D., Hamilton, E.P., Xiong, J., Lawson, M.J., Yuan, D., Hadjithomas, M., Miao, W., Orias, E., 2013. Selecting one of several mating types through gene segment joining and deletion in *Tetrahymena thermophila*. PLoS Biol. 11, 1–15.

Cervantes, M.D., Xi, X., Vermaak, D., Yao, M.-C., Malik, H.S., 2006. The CNA1 histone of the ciliate *Tetrahymena thermophila* is essential for chromosome segregation in the germline micronucleus. Mol. Biol. Cell 17, 485–497.

Chalker, D.L., Yao, M.-C., 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, D.L., Yao, M.-C., 2001. Nongenic, bidirectional transcription precedes and may promote developmental DNA deletion in *Tetrahymena thermophila*. Genes Dev. 15, 1287–1298.

Chen, X., Bracht, J.R., Goldman, A.D., Dolzhenko, E., Clay, D.M., Swart, E.C., Perlman, D.H., Doak, T.G., Stuart, A., Amemiya, C.T., 2014. The architecture of a scrambled genome reveals massive levels of genomic rearrangement during development. Cell 158, 1187–1198.

Chen, X., Gao, S., Liu, Y., Wang, Y., Wang, Y., Song, W., 2016. Enzymatic and chemical mapping of nucleosome distribution in purified micro- and macronuclei of the ciliated model organism, *Tetrahymena thermophila*. Sci. China: Life Sci. 59, 909–919.

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.

Cheng, X., 1995. Structure and function of DNA methyltransferases. Annu. Rev. Biophys. Biomol. Struct. 24, 293–318.

Coyne, R.S., Chalker, a., Douglas, L., Yao, M.-C., 1996. Genome downsizing during ciliate development: nuclear division of labor through chromosome restructuring. Annu. Rev. Genet. 30, 557–578.

Coyne, R.S., Stover, N., Miao, W., 2012. Whole genome studies of *Tetrahymena*. Methods Cell Biol. 109, 53–81.

Cui, B., Gorovsky, M.A., 2006. Centromeric histone H3 is essential for vegetative cell division and for DNA elimination during conjugation in *Tetrahymena thermophila*. Mol. Cell. Biol. 26, 4499–4510.

Cui, B., Liu, Y., Gorovsky, M.A., 2006. Deposition and function of histone H3 variants in *Tetrahymena thermophila*. Mol. Cell. Biol. 26, 7719–7730.

Cummings, D.J., Tait, A., Goddard, J.M., 1974. Methylated bases in DNA from *Paramecium aurelia*. Biochim. Biophys. Acta Nucleic Acids Protein Synth. 374, 1–11.

Dini, F., Luporini, P., 1979. The multiple mating type system of the marine ciliate *Euplotes crassus* (Dujardin). Arch. Protistenk. 121, 238–245.

Doak, T.G., Witherspoon, D.J., Jahn, C.L., Herrick, G., 2003. Selection on the genes of *Euplotes crassus* Tec1 and Tec2 transposons: evolutionary appearance of a programmed frameshift in a Tec2 gene encoding a tyrosine family site-specific ecombinase. Eukaryot. Cell 2, 95–102.

Doerder, F., Deak, J., Lief, J., 1992. Rate of phenotypic assortment in *Tetrahymena thermophila*. Dev. Genet. 13, 126–132.

Doerder, F.P., Gates, M.A., Eberhardt, F.P., Arslanyolu, M., 1995. High frequency of sex and equal frequencies of mating types in natural populations of the ciliate *Tetrahymena thermophila*. Proc. Natl. Acad. Sci. U. S. A. 92, 8715–8718.

Duharcourt, S., Butler, A., Meyer, E., 1995. Epigenetic self-regulation of developmental excision of an internal eliminated sequence on *Paramecium tetraurelia*. Genes Dev. 9, 2065–2077.

Duharcourt, S., Keller, A.-M., Meyer, E., 1998. Homology-dependent maternal inhibition of developmental excision of Internal Eliminated Sequences in *Paramecium tetraurelia*. Mol. Cell. Biol. 18, 7075–7085.

Duharcourt, S., Lepère, G., Meyer, E., 2009. Developmental genome rearrangements in ciliates: a natural genomic subtraction mediated by non-coding transcripts. Trends Genet. 25, 344–350.

Eisen, J.A., Coyne, R.S., Wu, M., Wu, D., Thiagarajan, M., Wortman, J.R., Badger, J.H., Ren, Q., Amedeo, P., Jones, K.M., Tallon, L.J., Delcher, A.L., Salzberg, S.L., Silva, J.C., Haas, B.J., Majoros, W.H., Farzad, M., Carlton, J.M., Smith Jr., R.K., Garg, J., Pearlman, R.E., Karrer, K.M., Sun, L., Manning, G., Elde, N.C., Turkewitz, A.P., Asai, D.J., Wilkes, D.E., Wang, Y., Cai, H., Collins, K., Stewart, B.A., Lee, S.R., Wilamowska, K., Weinberg, Z., Ruzzo, W.L., Wloga, D., Gaertig, J., Frankel, J., Tsao, C.C., Gorovsky, M.A., Keeling, P.J., Waller, R.F., Patron, N.J., Cherry, J.M., Stover, N.A., Krieger, C.J., del Toro, C., Ryder, H.F., Williamson, S.C., Barbeau, R.A., Hamilton, E.P., Orias, E., 2006. Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. PLoS Biol. 4, e286.

Fang, W., Wang, X., Bracht, J.R., Nowacki, M., Landweber, L.F., 2012. Piwi-interacting RNAs protect DNA against loss during *Oxytricha* genome rearrangement. Cell 151, 1243–1255.

Fass, J.N., Joshi, N.A., Couvillion, M.T., Bowen, J., Gorovsky, M.A., Hamilton, E.P., Orias, E., Hong, K., Coyne, R.S., Eisen, J.A., Chalker, D.L., Lin, D., Collins, K., 2011. Genome-scale analysis of programmed DNA elimination sites in *Tetrahymena thermophila*. G3: Genes, Genomes, Genet. 1, 515–522.

Feng, S., Jacobsen, S.E., Reik, W., 2010. Epigenetic reprogramming in plant and animal development. Science 330, 622–627.

Foissner, W., 1995. Tropical protozoan diversity: 80 ciliate species (Protozoa, Ciliophora) in a soil sample from a tropical dry forest of Costa Rica, with descriptions of four new genera and seven new species. Arch. Protistenk. 145, 37–79.

Foissner, W., Berger, H., 1999. Identification and ontogenesis of the nomen nudum hypotrichs (Protozoa: Ciliophora) *Oxytricha nova*

(=*Sterkiella nova* sp. n.) and *O. trifallax* (=*S. histriomuscorum*). Acta Protozool. 38, 215–248.

Foissner, W., Berger, H., Kohmann, F., 1992. Taxonomische und ökologische Revision der Ciliaten des Saprobiensystems—Band II: Peritrichia, Heterotrichida, Odontostomatida. Bayerisches Landesamt für Wasserwirtschaft, Informationsberichte des Bayer.

Fu, Y., Luo, G.-Z., Chen, K., Deng, X., Yu, M., Han, D., Hao, Z., Liu, J., Lu, X., Doré, L.C., 2015. N^6-methyldeoxyadenosine marks active transcription start sites in *Chlamydomonas*. Cell 161, 879–892.

Galvani, A., Sperling, L., 2002. RNA interference by feeding in *Paramecium*. Trends Genet. 18, 11–12.

Gao, F., Warren, A., Zhang, Q., Gong, J., Miao, M., Sun, P., Xu, D., Huang, J., Yi, Z., Song, W., 2016. The all-data-based evolutionary hypothesis of ciliated protists with a revised classification of the phylum Ciliophora (Eukaryota, Alveolata). Sci. Rep. 6, 24874.

Gorovsky, M., Glover, C., Johmann, C., Keevert, J., Mathis, D., Samuelson, M., 1978. Histones and chromatin structure in *Tetrahymena* macro- and micronuclei. Cold Spring Harb. Symp. Quant. Biol. 42, 493–503.

Gorovsky, M.A., Hattman, S., Pleger, G.L., 1973. [^{6}N] Methyl adenine in the nuclear DNA of a eucaryote, *Tetrahymena pyriformis*. J. Cell Biol 56, 697–701.

Greer, E.L., Blanco, M.A., Gu, L., Sendinc, E., Liu, J., Aristizábal-Corrales, D., Hsu, C.-H., Aravind, L., He, C., Shi, Y., 2015. DNA methylation on N^6-adenine in *Celegans*. Cell 161, 868–878.

Greslin, A.F., Prescott, D.M., Oka, Y., Loukin, S.H., Chappell, J.C., 1989. Reordering of nine exons is necessary to form a functional actin gene in *Oxytricha nova*. Proc. Natl. Acad. Sci. U. S. A. 86, 6264–6268.

Hamilton, E.P., Kapusta, A., Huvos, P.E., Bidwell, S.L., Zafar, N., Tang, H., Hadjithomas, M., Krishnakumar, V., Badger, J.H., Caler, E.V., 2016. Structure of the germline genome of *Tetrahymena thermophila* and relationship to the massively rearranged somatic genome. Elife 5, e19090.

Hammerschmidt, B., Schlegel, M., Lynn, D.H., Leipe, D.D., Sogin, M.L., Raikov, I.B., 1996. Insights into the evolution of nuclear dualism in the ciliates revealed by phylogenetic analysis of rRNA sequences. J. Eukaryot. Microbiol. 43, 225–230.

Harrison, G.S., Findly, R.C., Karrer, K., 1986. Site-specific methylation of adenine in the nuclear genome of a eucaryote, *Tetrahymena thermophila*. Mol. Cell. Biol. 6, 2364–2370.

Hattman, S., Kenny, C., Berger, L., Pratt, K., 1978. Comparative study of DNA methylation in three unicellular eucaryotes. J. Bacteriol. 135, 1156–1157.

Hausmann, K., Bradbury, P.C., 1996. Ciliates: Cells as Organisms. Spektrum Akademischer Verlag.

Hayashi, T., Hayashi, H., Koichi, I., 1987. *Tetrahymena* histone H1. isolation and amino acid sequence lacking the central hydrophobic domain conserved in other H1 histones. J. Biochem. 102, 369–376.

Henikoff, S., Smith, M.M., 2015. Histone variants and epigenetics. Cold Spring Harb. Perspect. Biol. 7, a019364.

Hiwatashi, K., 1968. Determination and inheritance of mating type in *Paramecium caudatum*. Genetics 58, 373.

Holliday, R., Pugh, J.E., 1975. DNA modification mechanisms and gene activity during development. Science 187, 226–232.

Jahn, C.L., Doktor, S.Z., Frels, J.S., Jaraczewski, J.W., Krikau, M.F., 1993. Structures of the *Euplotes crassus* Tec1 and Tec2 elements: identification of putative transposase coding regions. Gene 133, 71–78.

Jahn, C.L., Klobutcher, L.A., 2002. Genome remodeling in ciliated protozoa. Annu. Rev. Microbiol. 56, 489–520.

Juranek, S., Wieden, H.-J., Lipps, H.J., 2003. *De novo* cytosine methylation in the differentiating macronucleus of the stichotrichous ciliate *Stylonychia lemnae*. Nucleic Acids Res. 31, 1387–1391.

Kahl, A., 1931. Urtiere oder Protozoa I: Wimpertiere oder Ciliata (Infusoria). 2. Holotricha. Tierwelt. Dtl. 21, 181–398.

Karrer, K.M., VanNuland, T.A., 2002. Methylation of adenine in the nuclear DNA of *Tetrahymena* is internucleosomal and independent of histone H1. Nucleic Acids Res. 30, 1364–1370.

Kay, P., Pereira, E., Marlow, S., Turbett, G., Mitchell, C., Jacobsen, P., Holliday, R., Papadimitriou, J., 1994. Evidence for adenine methylation within the mouse myogenic gene Myo-D1. Gene 151, 89–95.

Klobutcher, L.A., 1999. Characterization of in vivo developmental chromosome fragmentation intermediates in *E crassus*. Mol. Cell 4, 695–704.

Koziol, M.J., Bradshaw, C.R., Allen, G.E., Costa, A.S., Frezza, C., Gurdon, J.B., 2016. Identification of methylated deoxyadenosines in vertebrates reveals diversity in DNA modifications. Nat. Struct. Mol. Biol. 23, 24–30.

Kwok, F., Ng, S.F., 1989. 5-Azacytidine affects the programming of expression of the somatic nucleus of *Paramecium*. Development 105, 559–568.

Lepère, G., Nowacki, M., Serrano, V., Gout, J.F., Guglielmi, G., Duharcourt, S., Meyer, E., 2009. Silencing-associated and meiosis-specific small RNA pathways in *Paramecium tetraurelia*. Nucleic Acids Res. 37, 903–915.

Lhuillier-Akakpo, M., Frapporti, A., Wilkes, C.D., Matelot, M., Vervoort, M., Sperling, L., Duharcourt, S., 2014. Local effect of enhancer of zeste-like reveals cooperation of epigenetic and *cis*-acting determinants for zygotic genome rearrangements. PLoS Genet. 10, 1–17.

Lin, C.-Y.G., Lin, I.-T., Yao, M.-C., 2016. Programmed minichromosome elimination as a mechanism for somatic genome reduction in *Tetrahymena thermophila*. PLoS Genet. 12, e1006403.

Liu, J., Zhu, Y., Luo, G.-Z., Wang, X., Yue, Y., Wang, X., Zong, X., Chen, K., Hang, Y., Fu, Y., Han, D., Wang, Y., Chen, D., He, C., 2016. Abundant DNA 6mA methylation during early embryogenesis of zebrafish and pig. Nat. Commun. 7, 13052.

Liu, Y., Taverna, S.D., Muratore, T.L., Shabanowitz, J., Hunt, D.F., Allis, C.D., 2007. RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in *Tetrahymena*. Genes Dev. 21, 1530–1545.

Lynn, D., 2008. The Ciliated Protozoa: Characterization, Classification, and Guide to the Literature, third ed. Springer Dordrecht.

Lynn, D.H., Small, E.B., 2002. Phylum Ciliophora Doflein, 1901. In: Lee, J.J., Leedale, G.G., Bradbury, P.C. (Eds.), The Illustrated Guide to the Protozoa. , second ed. Allen Press Inc., Lawrence, Kansas, pp. 371–656.

Malone, C.D., Anderson, A.M., Motl, J.A., Rexer, C.H., Chalker, D.L., 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.

Mochizuki, K., Fine, N.A., Fujisawa, T., Gorovsky, M.A., 2002. Analysis of a *piwi*-related gene implicates small RNAs in genome rearrangement in *Tetrahymena*. Cell 110, 689–699.

Mochizuki, K., Gorovsky, M.A., 2005. A Dicer-like protein in *Tetrahymena* has distinct functions in genome rearrangement, chromosome segregation, and meiotic prophase. Genes Dev. 19, 77–89.

Nanney, D.L., Caughey, P.A., 1953. Mating type determination in *Tetrahymena pyriformis*. Proc. Natl. Acad. Sci. U. S. A. 39, 1057–1063.

Noto, T., Kataoka, K., Suhren, J.H., Hayashi, A., Woolcock, K.J., Gorovsky, M.A., Mochizuki, K., 2015. Small-RNA-mediated genome-wide trans-recognition network in *Tetrahymena* DNA elimination. Mol. Cell 59, 229–242.

Nowacki, M., Higgins, B.P., Maquilan, G.M., Swart, E.C., Doak, T.G., Landweber, L.F., 2009. A functional role for transposases in a large eukaryotic genome. Science 324, 935–938.

Nowacki, M., Vijayan, V., Zhou, Y., Schotanus, K., Doak, T.G., Landweber, L.F., 2008. RNA-mediated epigenetic programming of a genome-rearrangement pathway. Nature 451, 153–158.

Orias, E., Cervantes, M.D., Hamilton, E.P., 2011. Tetrahymena thermophila, a unicellular eukaryote with separate germline and somatic genomes. Res. Microbiol. 162, 578–586.

Palacios, G., Martin-Gonzalez, A., Gutierrez, J.C., 1994. Macronuclear DNA demethylation is involved in the encystment process of the ciliate *Colpoda inflata*. Cell Biol. Int. 18, 201–206.

Parfrey, L.W., Lahr, D.J., Knoll, A.H., Katz, L.A., 2011. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. Proc. Natl. Acad. Sci. U. S. A. 108, 13624–13629.

Prescott, D.M., 1994. The DNA of ciliated protozoa. Microbiol. Rev. 58, 233–267.

Prescott, D.M., 1999. The evolutionary scrambling and developmental unscrambling of germline genes in hypotrichous ciliates. Nucleic Acids Res. 27, 1243–1250.

Prescott, D.M., Prescott, J.D., Prescott, R.M., 2002. Coding properties of macronuclear DNA molecules in *Sterkiella nova* (*Oxytricha nova*). Protist 153, 71–77.

Rae, P.M., Spear, B.B., 1978. Macronuclear DNA of the hypotrichous ciliate *Oxytricha fallax*. Proc. Natl. Acad. Sci. U. S. A. 75, 4992–4996.

Ratel, D., Ravanat, J.L., Berger, F., Wion, D., 2006. N6-methyladenine: the other methylated base of DNA. BioEssays 28, 309–315.

Ray, C., 1956. Meiosis and nuclear behavior in *Tetrahymena pyriformis*. J. Protozool. 3, 88–96.

Razin, A., Cedar, H., 1977. Distribution of 5-methylcytosine in chromatin. Proc. Natl. Acad. Sci. 74, 2725–2728.

Reyes, E.M., Camacho-Arroyo, I., Nava, G., Cerbon, M.A., 1997. Differential methylation in steroid 5α-reductase isozyme genes in epidydimis, testis, and liver of the adult rat. J. Androl. 18, 372–377.

Rogers, M.B., Karrer, K.M., 1985. Adolescence in *Tetrahymena thermophila*. Proc. Natl. Acad. Sci. U. S. A. 82, 436–439.

Russo, V.E., Martienssen, R.A., Riggs, A.D., 1996. Epigenetic Mechanisms of Gene Regulation. Cold Spring Harbor Laboratory Press, New York.

Salim, H.M., Ring, K.L., Cavalcanti, A.R., 2008. Patterns of codon usage in two ciliates that reassign the genetic code: *Tetrahymena thermophila* and *Paramecium tetraurelia*. Protist 159, 283–298.

Salvini, M., Barone, E., Ronca, S., Nobili, R., 1986. DNA methylation in vegetative and conjugating cells of a protozoan ciliate: *Blepharisma japonicum*. Dev. Genet. 7, 149–158.

Salvini, M., Durante, M., Citti, L., Nobili, R., 1984. 5′-Methylcytosine in the macronulear DNA of *Blepharisma japonicum*. Experientia 40, 1401–1403.

Sandoval, P.Y., Swart, E.C., Arambasic, M., Nowacki, M., 2014. Functional diversification of Dicer-like proteins and small RNAs required for genome sculpting. Dev. Cell 28, 174–188.

Schoeberl, U.E., Kurth, H.M., 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.

Shieh, A.W., Chalker, D.L., 2013. LIA5 is required for nuclear reorganization and programmed DNA rearrangements occurring during *Tetrahymena* macronuclear differentiation. PLoS One 8, e75337.

Siegel, R., 1956. Mating types in *Oxytricha* and the significance of mating type systems in ciliates. Biol. Bull. 110, 352–357.

Singh, D.P., Saudemont, B., Guglielmi, G., Arnaiz, O., GoÛt, J.-F., Prajer, M., Potekhin, A., Przybòs, E., Aubusson-Fleury, A., Bhullar, S., 2014. Genome-defence small RNAs exapted for epigenetic mating-type inheritance. Nature 509, 447–452.

Slabodnick, M.M., Ruby, J.G., Reiff, S.B., Swart, E.C., Gosai, S., Prabakaran, S., Witkowska, E., Larue, G.E., Fisher, S., Freeman, R.M., 2017. The macronuclear genome of *Stentor coeruleus* reveals tiny introns in a giant cell. Curr. Biol. 27, 569–575.

Smith, Z.D., Meissner, A., 2013. DNA methylation: roles in mammalian development. Nat. Rev. Genet. 14, 204–220.

Song, W., Warren, A., Hu, X., 2009. Free-living Ciliates in the Bohai and Yellow Seas, China, first ed. Science Press, Beijing.

Sonneborn, T.M., 1937. Sex, sex inheritance and sex determination in *Paramecium aurelia*. Proc. Natl. Acad. Sci. U. S. A. 23, 378–385.

Strahl, B.D., Ohba, R., Cook, R.G., Allis, C.D., 1999. Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in *Tetrahymena*. Proc. Natl. Acad. Sci. U. S. A. 96, 14967–14972.

Swart, E.C., Bracht, J.R., Magrini, V., Minx, P., Chen, X., Zhou, Y., Khurana, J.S., Goldman, A.D., Nowacki, M., Schotanus, K., 2013. The *Oxytricha trifallax* macronuclear genome: a complex eukaryotic genome with 16,000 tiny chromosomes. PLoS Biol. 11, e1001473.

Swart, E.C., Serra, V., Petroni, G., Nowacki, M., 2016. Genetic codes with no dedicated stop codon: context-dependent translation termination. Cell 166, 691–702.

Tajima, S., Suetake, I., 1998. Regulation and function of DNA methylation in vertebrates. J. Biochem. 123, 993–999.

Taverna, S.D., Coyne, R.S., Allis, C.D., 2002. Methylation of histone H3 at lysine 9 targets programmed DNA elimination in *Tetrahymena*. Cell 110, 701–711.

Taylor, S., Constantinides, P., Jones, P., 1983. 5-Azacytidine, DNA methylation, and differentiation. Curr. Top. Microbiol. Immunol. 108, 115–127.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., Martienssen, R.A., 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science 297, 1833–1837.

Wang, R., Xiong, J., Wang, W., Miao, W., Liang, A., 2016. High frequency of +1 programmed ribosomal frameshifting in *Euplotes octocarinatus*. Sci. Rep. 6, 21139.

Wion, D., Casadesús, J., 2006. N^6-methyl-adenine: an epigenetic signal for DNA–protein interactions. Nat. Rev. Microbiol. 4, 183–192.

Wu, M., Allis, C.D., Richman, R., Cook, R.G., Gorovsky, M.A., 1986. An intervening sequence in an unusual histone H1 gene of *Tetrahymena thermophila*. Proc. Natl. Acad. Sci. U. S. A. 83, 8674–8678.

Wu, T.P., Wang, T., Seetin, M.G., Lai, Y., Zhu, S., Lin, K., Liu, Y., Byrum, S.D., Mackintosh, S.G., Zhong, M., 2016. DNA methylation on N^6-adenine in mammalian embryonic stem cells. Nature 532, 329–333.

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

Yao, M.-C., Gall, J.G., 1977. A single integrated gene for ribosomal RNA in a eucaryote, *Tetrahymena pyriformis*. Cell 12, 121–132.

Yao, M.-C., Gorovsky, M.A., 1974. Comparison of the sequences of macro- and micronuclear DNA of *Tetrahymena pyriformis*. Chromosoma 48, 1–18.

Yao, M.-C., Yao, C.-H., Monks, B., 1990. The controlling sequence for site-specific chromosome breakage in *Tetrahymena*. Cell 63, 763–772.

Zhang, G., Huang, H., Liu, D., Cheng, Y., Liu, X., Zhang, W., Yin, R., Zhang, D., Zhang, P., Liu, J., 2015. N^6-methyladenine DNA modification in *Drosophila*. Cell 161, 893–906.

Zhao, X., Wang, Y., Wang, Y., Liu, Y., Gao, S., 2016. Histone methyltransferase TXR1 is required for both H3 and H3. 3 lysine 27 methylation in the well-known ciliated protist *Tetrahymena thermophila*. Sci. China. Life Sci. 59, 1–7.

Zoller, S.D., Hammersmith, R.L., Swart, E.C., Higgins, B.P., Doak, T.G., Herrick, G., Landweber, L.F., 2012. Characterization and taxonomic validity of the ciliate *Oxytricha trifallax* (Class Spirotrichea) based on multiple gene sequences: limitations in identifying genera solely by morphology. Protist 163, 643–657.