Programmed genome rearrangements in ciliates

Iwona Rzeszutek1 · Xyrus X. Maurer-Alcalá2 · Mariusz Nowacki2

Received: 6 January 2020 / Revised: 11 May 2020 / Accepted: 15 May 2020 / Published online: 27 May 2020
© The Author(s) 2020

Abstract
Ciliates are a highly divergent group of unicellular eukaryotes with separate somatic and germline genomes found in distinct dimorphic nuclei. This characteristic feature is tightly linked to extremely laborious developmentally regulated genome rearrangements in the development of a new somatic genome/nuclei following sex. The transformation from germline to soma genome involves massive DNA elimination mediated by non-coding RNAs, chromosome fragmentation, as well as DNA amplification. In this review, we discuss the similarities and differences in the genome reorganization processes of the model ciliates Paramecium and Tetrahymena (class Oligohymenophorea), and the distantly related Euplotes, Stylonychia, and Oxytricha (class Spirotrichea).

Keywords Ciliates · Nuclear dimorphism · Genome rearrangement · DNA elimination · Small non-coding RNAs

Introduction
Developmentally regulated genome rearrangements (DRGRs) involve the elimination of specific DNA sequences (from the germline) somatic cell lineages. In most cases, this phenomenon is associated with two forms of DNA elimination either: (a) chromosome elimination where the entire chromosome is lost [1] or (b) chromosome diminution, a process characterized by loss of chromosome portions through chromosome breakage and repair during the developmental transformation from germline to soma [2, 3].

Programmed DNA elimination was first described in 1887 by Theodor Boveri [4] in the horse parasitic nematode, Parascaris univalens. Since then, DRGRs have been identified in diverse multicellular organisms including nematodes, arthropods, hagfish, lampreys [3] and lymphoid lineages of vertebrates [5]. However, it appears most pervasive in ciliates, an ancient clade of microbial eukaryotes (> 1 Gya; [6]), where genome rearrangements lead to the elimination of 30–95% of the germline genome [7–9]. This review will focus on genome rearrangements in the two best-studied classes of ciliates: the Oligohymenophorea (including Paramecium and Tetrahymena), and members of the Spirotrichea (including Euplotes, Oxytricha and Stylonychia).

Ciliates as a model organism
Ciliates are unicellular eukaryotes found in diverse environments (fresh/saltwater as well as soil) across the globe that emerged more than 1 billion years ago [6]. Due to their morphological and morphogenetic characters, the taxonomy of ciliates has been ambiguous for a long time. Numerous studies have improved the phylogenetic relationship between ciliates with the rest of the eukaryotic tree of life, being members of the Alveolata (along with apicomplexans and dinoflagellates) [10]. Similarly, phylogomic studies within Ciliophora illustrate the great diversity and deep evolutionary history despite limited taxon sampling [11]. Even though just a handful of the ~ 4500 described ciliate species have been studied in-depth [12], they share complex cytoskeletal structures, well-developed ciliary structures at the cell surface (for swimming, food uptake and sensing environmental signals), the separation of germline and somatic genomes into distinct nuclei (nuclear dimorphism), as well as DRGRs (reviewed in [13]). Although the majority of studies are

1 Institute of Biology and Biotechnology, Department of Biotechnology, University of Rzeszow, Pigonia 1, 35-310 Rzeszow, Poland
2 Institute of Cell Biology, University of Bern, Baltzerstrasse 4, 3012 Bern, Switzerland
limited to the model genera *Tetrahymena*, *Paramecium*, and *Oxytricha*, these models have greatly contributed to our understanding of biological mechanisms and phenomena present in diverse eukaryotic lineages. This includes discovery and description of ribozymes [14], the discovery of the first histone-modifying enzyme [15], variant nuclear genetic codes [16–20], the initial identification of telomerase and telomere structure [21, 22], numerous examples of small RNA-mediated heterochromatin formation [23, 24], as well as mechanisms enabling the transcription of short DNA fragments [25].

**Nuclear dimorphism in ciliates**

In multicellular eukaryotes, germline and somatic functions are separated into distinct cell types, (e.g. pollen versus leaf in plants, spor versus hyphae in fungi, or egg versus skin in humans). However, in ciliates, both germline and somatic genomes co-exist within a single cell, providing each cell with at least one somatic nucleus used for gene expression and one germline to propagate the genome across sexual generations (Fig. 1).

Each ciliate cell possesses at least one micronucleus (MIC) and one macronucleus (MAC); however, their number varies between the species (reviewed in [26]). Interestingly, all micronuclei present in the cell possess features of typical eukaryotic nuclei, (i.e. diploid [27–29], centromeres [30] and are transposon rich [30, 31]). The MAC is transcriptionally active throughout the entire life cycle and possesses highly processed chromosomes. These MAC chromosomes are gene-rich, lack centromeres and can range in ploidy from ~2 N in the Karyorelictea to > 13,000 N in the Heterotricha [32–34]. During asexual growth, the hyper-polyploid MACs divide amitotically, which lacks mitotic spindles and chromatin condensation, separating chromosomes in bulk as large masses, which can result in daughter nuclei with unequal amounts of DNA. The degree of inequality in the segregation of DNA to the two MACs during amitosis can be exacerbated under environmental stressors [35, 36]. Moreover, MAC chromosomes are amplified to elevated copy numbers [26]. For instance, in Oligohymenophorea, each MAC chromosome is present in the equal copy number, namely, ~ 45 copies of each 225 chromosomes in *T. thermophila* [37, 38] and ~ 800 copies of each chromosome in *P. tetraurelia* [39, 40]. In contrast, the *Oxytricha* MAC harbours thousands of unique gene-sized nanochromosomes, which are amplified to ~ 1900 copies [31]. However, unlike *Paramecium* and *Tetrahymena*, these nanochromosomes are maintained at unique copy numbers, varying between a few hundred to 10^6 copies [26, 41–43].

During vegetative (or asexual) growth, the germline remains transcriptionally inactive and divides mitotically [44]. This changes during sex or self-fertilisation/autogamy (Fig. 2) [45]. At the onset of sex and development, micronuclei undergo meiosis and are fused with a partner haploid MIC that gives rise to the zygotic nucleus from which new micro- and macronuclei are formed.

**Features and origin of internal eliminated sequences**

The discovery that germline-limited internal eliminated sequences (IESs) resemble transposon sequences brought a new challenge for scientists trying to understand the mechanisms underlying their elimination. Jacobs and Kloobutcher [46] observed that IESs in *Euplotes crassus* possess the consensus sequence 5′-TATrGCRN-3′ (Y = pyrimidine, R = purine), which resembles terminal inverted repeats (TIRs) at the end of their Tec family transposable elements.
Programmed genome rearrangements in ciliates

[46–48] and Paramecium’s Tcl/Mariner transposons [46–49]. All IESs found to date in Euplotes and Paramecium possess 5′-TA-3′ dinucleotide repeats at each boundary, where a single copy of the dinucleotide remains in the new macronucleus after excision. Based on these observations, Klobutcher and Herrick [47, 50] developed a model for the origin of IESs, where transposons initially invade the germline genome, then spread throughout, and ultimately decay over time into the identifiable IESs currently found in ciliate germline genomes.

These sequences must be excised during the development of a MAC to produce a functional somatic genome. IESs are present in all the ciliate germline genomes studied to date [30, 31, 48, 51–53], although in varying amounts (~12,000 IESs in Tetrahymena thermophila [54], ~45,000 IES in Paramecium tetraurelia [48, 51] and >200,000 IESs in Oxytricha trifallax [31]. IESs are typically AT rich (70–100%) and bounded by pairs of short direct repeats (most are 1-8 bp) that help identify the boundaries between Macronuclear Destined Sequences (MDSs) and IESs [30, 31, 48, 52]. Recently, Maurer-Alcalá et al. [53] demonstrated that MDS–IES boundaries are identifiable by sharp changes in GC content. For instance, in P. tetraurelia where IES excision is precise, GC content was decreased in close proximity to its MDS–IES boundaries. On the other hand, in Tetrahymena where almost all IESs are excised imprecisely, GC contents are characterized by the great variability associated with MDS–IES boundaries within the inferred MDS itself [53]. The length and genomic distribution of IESs in germline genomes are very diverse, with most IESs in Tetrahymena being intergenic and “long” (>100 bp to over 10 Kbp) [55, 56], whereas the IESs in Paramecium and Oxytricha often interrupt protein-coding sequences and are comparatively short (most <100 bp) [31, 48].

As IESs often disrupt coding regions for most ciliates, they must be accurately excised during development to enable expression of the functional genes in the newly developed MAC [48, 57–59]. However, in Tetrahymena...
**thermophila** nearly all of the 7350 well-described IESs (of the ~12,000 total IESs) are excised imprecisely, due to variable MDS–IES junction sites. The potential deleterious impacts of this imprecision are likely mitigated by the genomic distribution of these IESs, which are predominantly found in intergenic (6182; 82%) or intronic (1168; 16%) regions [30, 60–62]. In contrast, only the excision of transposon-like sequences and minisatellites in *Paramecium* is imprecise [63], whereas the ~45,000 IESs nestled within or near protein-coding sequences are precisely excised during macronuclear development, although small numbers of IESs are excised at alternative MDS–IES boundaries [40, 48, 64]. Overall, imprecise elimination results either in the fragmentation of micronuclear chromosomes into shorter acentromeric macronuclear chromosomes to which telomeric repeats are added, or to the imprecise re-joining of flanking sequences [63].

Previous work in *Paramecium* and *Tetrahymena* have shown that the pointer sequences present at both ends of IESs, influences the efficacy of IES excision. Analyses of *Paramecium*’s IESs have demonstrated that single base-pair-mutations in the conserved terminal repeat of IESs lead to their retention during development [65–67]. Additionally, for some *Paramecium* IESs, flanking sequences are necessary for excision [68]. For example, the removal of a portion of the 72 bp flanking region of one end of a small 28 bp IES in *Paramecium* reduced the efficiency of excision, and complete removal of all wild-type sequences adjacent to the TA abolished excision [68]. Recently, it has been shown a small subset of *Paramecium*’s IESs shares a common 5 bp motif that is implicated in their sRNA-independent excision [69]. In *Tetrahymena*, flanking sequences are known to have a significant role in the elimination of a number of IESs [70, 71]. Together, Lia3p and Lia3-Like 1 (LTL1) regulatory proteins interact with flanking regulatory sequences to determine MDS–IES boundaries for several IESs for excision [72–74]. These data highlight the importance of both IES pointer sequences and their flanking regions in identifying MDS–IES boundaries.

Another variable that influences IES excision/recognition in *Paramecium* is its length. Swart et al. [75] indicated that the frequencies of IES sub-terminal bases change with IES length. Moreover, it has been shown that small IESs (shorter than 150 bp) are less sensitive to sRNAs depletion [76], suggesting that some IESs are more difficult to recognize/excise and require additional information (from the sRNAs) for their accurate excision.

### Developmentally regulated genome rearrangements

Although the mechanistic details behind ciliate DRGRs differ between even closely related taxa (*i.e.* *P. tetraurelia* and *T. thermophila*), the basic principles of this phenomenon are conserved (Table 1). In the developing macronucleus, rapid DNA synthesis takes place and interstitial DNA sequences such as transposons, minisatellites and IESs are excised. Afterward, the hundreds to thousands of broken chromosome ends created during excision are rejoined through non-homologous end-joining mechanisms [77–79], followed by de novo telomere addition, and finally chromosome amplification (Fig. 3). The end result is the production of a new functional somatic nucleus that contains the streamlined transcriptionally active chromosomes that maintain cell. This section is devoted to describing these phenomena in more detail.

### Role of small noncoding RNA in programmed genome rearrangements

An important breakthrough in our understanding of the regulation of DNA elimination was unravelling the involvement of noncoding RNA (ncRNA) in this complex epigenetic process. Briefly, the MIC is bi-directionally transcribed

---

**Table 1** Differences between the genome reorganization processes among ciliates

|                     | *Paramecium tetraurelia* | *Tetrahymena thermophila* | *Oxytricha trifallax* |
|---------------------|--------------------------|---------------------------|-----------------------|
| MIC chromosomes     | ?                        | 5 [37]                    | ?                     |
| MAC chromosomes     | 188 [39, 40]             | 225 [38]                  | 15,600 [43]           |
| Nanochromosomes     | No                       | No                        | Yes [43]              |
| Unscrambling        | No                       | No                        | No*a                  |
| IES percentage      | 30% [48]                 | 30% [179]                 | 90% [31]              |
| IES location        | Genic and intergenic regions [48] | Intergenic Regions [90] | Genic and intergenic regions [180] |
| Small RNA source    | MDS and IES [25, 81]     | Biased towards IES [111]  | MDS [83]              |
| Small RNA target    | IES [81]                 | IES [80]                  | MDS [82, 83]          |
| DNA methylation     | 6 mA (*P. aurelia*) [132] | 6 mA [131]                | 5mC, 6 mA [127, 135]  |

*aOne locus showing unscrambling [181]*
producing long transcripts that are processed into small RNAs. In Paramecium and Tetrahymena (cl: Oligohymenophorea), these sRNAs then “scan” the parental MAC. Those sRNAs corresponding to the parental MAC are lost, leading to the enrichment of MIC-matching sRNAs [23, 76, 80, 81]. However, in Oxytricha and Stylonychia these sRNAs are produced from transcripts derived from the parental MAC genome and are putatively involved in protecting MDSs in the MIC rather than identifying IESs for elimination (as in Paramecium and Tetrahymena) [82–84]. This reflects the extreme differences in germline genome content as ~ 30% of the germline genome is eliminated in Oligohymenophorean ciliates versus ≥ 95% of the germline in spirotrichs. These sRNAs pools ultimately help delineate MDSs and IESs, although details differ among even “closely” related species (e.g. Paramecium and Tetrahymena).

Small RNA-mediated programmed genome rearrangements in Oligohymenophorea

The first insights that DNA elimination relies on homologous RNA molecules originated from work in Oligohymenophorean ciliates (i.e. Paramecium and Tetrahymena). During prophase of meiosis, the MIC is bi-directionally transcribed [85] by RNA polymerase II [86], generating large MIC-based transcripts. These transcripts are then processed into small “scan” RNAs (sRNAs are 25 nt in Paramecium and 25–29 nt in Tetrahymena), by Dicer-like ribonuclease (Dcl1 in Tetrahymena and Dcl2 and Dcl3 in Paramecium), that “scan” for homologous sequences in the parental MAC genome (Fig. 4a) [76, 81, 87–89]. Initial experiments suggested that sRNA production from MIC-derived transcripts, represented relatively equal representation of IES and MDS regions [23]. However, recent work in Tetrahymena shows two “pulses” of sRNA production, which are associated with enriched transcription and processing of IES sequences [90]. Interestingly, the biased production of Tetrahymena sRNAs is predominantly from type-A IESs (per-centromeric and telomeric regions of the MIC chromosomes) [91].

After production in the micronuclei, sRNA duplexes are transported to the cytoplasm where they are loaded onto the PIWI family proteins, (Twi1p in Tetrahymena [23, 80] or Piwi1/9 complex present in Paramecium [92, 93]) which are then transported to the parental macronucleus (Fig. 4b). There, the genome “scanning” effectively removes the MAC-matching sRNAs, enriching for micronuclear-limited sRNAs from the initial population. These MIC-enriched sRNAs are transported to the developing macronucleus to guide DNA excision (Fig. 4c). Interestingly, it has been suggested that Paramecium’s sRNAs bind to longer RNA transcripts, rather than directly to DNA, in both the old and new MAC [94, 95]. In both Tetrahymena and Paramecium, there is a second wave of sRNAs that aid in ensuring the accurate identification and excision of IESs [76, 91]. In Tetrahymena, these “late” sRNAs are produced from both types of IESs [Type-A and Type-B (located at the chromosomal arms)] in cis. These late sRNAs are loaded onto the Twi1p and Tw11p complexes, which further guide heterochromatin formation in trans and ensure the elimination of all IESs copies [91]. However, in Paramecium, excised IESs are eventually circularized, with smaller IESs concatenated together prior to circularization [25]. These IES concatemers act as the transcriptional template for iesRNAs, which further ensure the precise and accurate excision of IESs [25]. These secondary iesRNAs are produced after IES excision in Paramecium, whereas Tetrahymena’s “late” sRNAs are produced prior to any IES excision [96]. Compared to
Tetrahymena where early and late scnRNAs are produced by the same Dicer-like ribonuclease (Dcl1), Paramecium iesRNAs are produced by Dcl5 [76]. As in Tetrahymena the primary scnRNAs and secondary iesRNAs are associated with distinct Piwi proteins, Ptiwi01/09 with scnRNAs and Ptiwi10/11 carries the iesRNAs [93]. Together, the primary and secondary sRNAs ensure the faithful elimination of all copies of IESs present in the developing MAC genome leading to the production of a new functional macronucleus (Fig. 4d).

Small RNA-mediated programmed genome rearrangements in Spirotrichea

DNA elimination in Oxytricha (Spirotrichea) is quite distinct from the distinctly related Oligohymenophorea. While Paramecium and Tetrahymena generate scnRNA in the parental MIC, Oxytricha’s, 27-nt-long small RNAs derive from the transcription of the parental macronucleus rather than the germline [82, 83]. In addition to that, these 27mers have been shown to associate with PIWI homologs called Otwi1 (hence called PIWI-interacting RNAs, piRNAs). The injection of 27 nt piRNAs corresponding to IESs into developing Oxytricha leads to the retention of those IES in the new somatic genome [83]. These data, combined with the apparent parental MAC origin of the piRNAs, suggest that they are responsible for identifying macronuclear destined sequences (MDS) to protect against excision, rather than targeting IESs for excision, as in Tetrahymena and Paramecium. Interestingly, as in Paramecium, Oxytricha also circularizes some excised TE and non-repetitive germline-limited sequences that are also actively transcribed [97, 98]. This presence of development-specific extrachromosomal circular DNA was originally described in Euplotes (cl: Spirotrichea) [99, 100], although the circularization process and content appears to differ between Oxytricha and Paramecium. If these circularized products of excised IES and transposon-like Tec elements in Euplotes lead to the production of small RNAs remains undetermined. Despite the differences in sRNA sources and targets, ciliates have evolved a relatively efficient and low-energy cost means to distinguish soma and germline.

DNA uns scrambling

In addition to delineating somatic and germline-limited DNA, macronuclear development in Oxytricha requires a very spectacular form of DNA rearrangement called uns crambling [26, 31, 101]. In the germline, MDSs can also be disordered and/or found on both strands of DNA (i.e. “inverted”) and may even originate from distant germline loci (Fig. 5) [31]. An extreme example is DNA polymerase α in O. nova, O. trifallax [102, 103] and S. lemnæ [104], which is broken into more than 40 MDS present at two distinct loci separated by > 3 kbp. Work in Oxytricha has demonstrated that DNA unscrabling is directed by long RNA templates derived from the parental macronucleus [105–107]. These long template RNAs, in conjunction with unique pointer sequences, act as a reference aiding in the accurate reordering of MDSs [108, 109]. The accuracy of
DNA unscrambling is incredibly sensitive to these template RNAs. For example, microinjection of alternately unscrambled templates (e.g. swapping the order of MDSs) leads to the production of macronuclear chromosome resembling the introduced template [105]. Furthermore, RNAi knockdown of these long RNA templates results in aberrant or reduced rearrangements of MDSs in the resulting chromosomes found in the new MAC [105].

**Histone modification in DNA elimination**

As in other eukaryotes, histone modifications play an integral role in the effective silencing of transposable elements and germline-limited DNA. In *Tetrahymena*, heterochromatin-specific marks, H3K9me3 and H3K27me3 are present in the MAC or both MIC and MAC, respectively [24, 87, 110–112]. Accumulation of H3K9me3 and H3K27me3 in the MAC is catalysed by histone methyltransferase Ezl1p, whereas Ezl2p is responsible for H3K27me3 in the MIC [111, 113]. As in other eukaryotes, small RNAs are involved in guiding the deposition of these conserved marks [110]. After deposition, specific marks are subsequently recognized by chromodomain-containing effectors. In particular, *Tetrahymena’s* Pdd1p (a homolog of HP1) accumulates on IESs, binding to methylated histones [114, 115] and is proposed to aid in recruiting *Tetrahymena’s* domesticated PiggyBac transposase (Tpb2p) for their excision [116]. Additionally, recent data indicated that RNAi-dependent *Polycomb* repression pathway is important for controlling transposable elements in *Tetrahymena* [117]. Disruption of the *Polycomb* repression pathway (knockout of DCL1, EZL1 and PDD1) results in the activation of TE transcription as well as the germline mobilization of TE [87, 110]. Moreover, numerous other histone modifications in *Tetrahymena* have been identified and may play roles in its DRGR [118, 119]. As in *Tetrahymena*, histone-specific marks such as H3K27 and H3K9 trimethylation are mediated by Ezl1 in *Paramecium* [120] and associated with chromatin assembly factor 1 subunit C-like protein (PtCAF-1) [121]. However, the developmental roles of these marks remain unclear.

As in Oligohymenophorea, heterochromatinization has been observed in Spirotrichea. During development, germline chromosomes are polytenized, with large blocks of observable heterochromatin prior to DNA elimination and fragmentation into thousands of unique gene-sized nanochromosomes [122]. In *Stylonychia*, this process is linked to the differential expression of a suite of histone H3 variants and subsequent post-translational modifications (PTM) [123–125]. For instance, H3K27me3 was shown to accumulate at the MIC-specific sequences prior to excision [126], while H3.7, acetylated at lysine-32, specifically associate with MDSs [123]. Moreover, knockdown of Piwi impacts the expression of histone H3.3 during macronuclear development in *Stylonychia*, implicating that H3.3 incorporation into nucleosomes is ncRNA-dependent [123]. Unfortunately, little is known about the roles of histone modifications and variants among other spirotrich ciliates, in *Oxytricha* and *Euplotes*.

**DNA modification in IES elimination**

Besides marking IESs for elimination through histone modifications, chemical modifications of germline DNA may also play a role in ciliate DRGRs. 5-Methylcytosine (5mC) has been identified on some germline limited sequences (e.g. transposons and satellite repeats) in *Stylonychia* and *Oxytricha* [127, 128] as well as in aberrant DNA rearrangements and parental DNA undergoing degradation [127].
Additionally, azacytidine and decitabine (DNA methyltransferase-inhibiting drugs) induce demethylation of both somatic and germline DNA during DRGRs, further implicating 5mC as a specific marker for DNA elimination/degradation [127]. Moreover, 5mC in *Stylochona* correlates with gene activity as well as with chromatin structure during macronuclear differentiation [129]. However, recent work in *Paramecium* was unable to detect any evidence for 5mC modifications, suggesting that these modifications may only be involved in the DRGRs of spirotrich ciliates [130].

While the function of 5mC DNA modifications and their phylogenetic distribution in ciliates remains unclear, the only widely conserved DNA modification is 6N-methyladenine (9 mA) [131–135]. Data from *Tetrahymena* shows that 6 mA is only present in the transcriptionally active MAC and is preferentially enriched in the consensus sequence 5′-AT-3′ [133, 134]. 6 mA modifications also localize to linker DNA regions downstream of the transcription start site (TSS of Polymerase II transcribed genes) and directly influence DNA regions downstream of the transcription start site (TSS). Tetrahymena’s MT-A70 homologue ATM1 (6mA DNA methyltransferase) responsible for 6mA modifications in *Oxytricha* disfavours nucleosome occupancy, contrary to *Tetrahymena* [135]. Beh et al. [135] suggest that decreased nucleosome occupancy is due to dA:dT base pair destabilisation by 6mA, which decreases the DNA melting temperature. However, the exact mechanism remains undetermined.

Unfortunately, while DNA modifications during MAC development are present, exactly how they might direct IES excision and/or MDS protection requires further investigation.

**Transposases required for DNA excision**

Transposase domestication has occurred throughout the eukaryotic tree of life, and can be linked to important DRGRs, such as those in ciliates and those involved in V(D) J recombination in animals [140, 141]. In *Paramecium* and *Tetrahymena*, IES excision is performed by a domesticated PiggyBac (PB) transposase (PiggyMac (Pgm) and TPB2 in *Paramecium* and *Tetrahymena*, respectively) [142, 143]. Excision of IESs in *Paramecium* by Pgm generates a 4 bp overhang at 5′ends centred around the “TA” dinucleotide pointer sequence showing the same geometry as those catalyzed in vitro by PB transposases [142, 144]. While Pgm is believed to carry out the physical excision, *Paramecium* possesses five accessory Pgm-like domesticated transposases (PgmL1-PgmL5) that interact with Pgm individually, with PgmL1 and PgmL3 directly involved in Pgm’s ability for precise IES excision [145]. Compared to *Paramecium’s* PiggyBac transposase, that possesses a sequence specificity (5′-TTAA-3′) [142], Tpb2p in *Tetrahymena* possess less stringent sequence specificity, as most of the IESs it excises are not flanked by any common motif [50, 62]. In *Tetrahymena* excision of the >10,000 Tpb2p-dependent IESs is imprecise [116, 143]. *Tetrahymena* also possesses multiple transposase proteins, such as Lia5p (domesticated PB transposase) which localizes on IESs and facilitates Tpb2p-dependent IES elimination [114, 146] and is required for chromosome fragmentation in *Tetrahymena* [146]. In contrast to the most abundant IESs in *Tetrahymena*, the excision of 12 particular IESs (possessing TE features, such as terminal inverted repeats and the 5′-TTAA-3′ cutting site) is precise and depends on Tpb1p and Tpb6p [147, 148].

In *Oxytricha*, IES excision is triggered by telomere-bearing element (TBE) family transposases [149], which belong to a superfamily of transposase genes that possess a common DDE catalytic motif [150]. Analyses of the *Oxytricha* genome found that complete and partial copies of *Tc1/ Mariner* transposons constitute around 13% of the germline genome [151]. Compared to the Paramecium’s and Tetrahymena’s transposase-related proteins, TBEs in *Oxytricha* are encoded in the MIC genome itself, rather than the somatic genome [152], and cut with a 3 nt 5′ overhang at an ANT recognition site [98]. Similar to members of the Oligohymenophorea, *Oxytricha* possesses multiple transposases, all of which have a necessary role in its development [149].

**Chromosome fragmentation**

Chromosome fragmentation is one of the major events that occur during macronuclear development in ciliates. The most extreme chromosome fragmentation takes place in ciliates with gene-sized chromosomes (e.g. *Euplotes, Stylochona*, and *Oxytricha*) where their MIC chromosomes are fragmented into >15,000 unique MAC chromosomes [43, 153]. Given the incredibly short size of these nanochromosomes (averaging ~2.8–3.2 Kbp in *Oxytricha* and *Stylochona* [43, 153]), most (~90%) encode just a single open reading frame (ORF) [43, 154]. While less dramatic, the five MIC chromosomes present in *T. thermophila* are fragmented into ~225 multigene MAC chromosomes (from <100–1500 Kbp) [155], whereas in *P. primaurelia*, MIC chromosome fragmentation gives rise to 50–1000 Kbp MAC chromosomes [156].

For some ciliates, the fragmentation of the MIC genome does not occur randomly, but at specific chromosome breakage sequences (CBS) in the germline [157]. In *Euplotes crassus*, a conserved 10 bp consensus sequence (*Euplotes*-chromosome breakage sequence; E-CBS: 5′-HATTGAA aHH’, H = A, C or T) directs a staggered double-strand break (DSB) at a precise distance and orientation, which provides the substrates for telomere addition [158–160]. In *Tetrahymena*, a conserved 15 bp chromosome breakage
sequence (CBS: 5′-WAAAACCAACCYCNHW-3′, W = A/T; Y = T/C; H = A/T/C; N = G/A/T/C) is necessary for chromosome fragmentation and telomere addition [30, 161–164]. While the E-CBSs in *Euplotes* are ultimately retained in the MAC [158–160], in *Tetrahymena* the CBSs themselves are germline-limited and are eliminated with 4–34 bp of flanking DNA on both sides [165]. Interestingly, in *Tetrahymena*, fragmentation of the germline at the CBSs generates 33 non-maintained chromosomes (NMCs) [30, 164, 166, 167]. Unlike the typical “large” (> 100 Kbp) somatic chromosomes in the MAC, these NMCs are generally short, ranging from 30 to 80 Kbp and have a limited life-span, either being degraded prior to de novo telomere addition or lost by ~120 asexual divisions [166]. While a majority of the NMCs harbour functional ORFs (some of which are actively transcribed) [30], it remains unclear what their role, if any, might be in the post-sexual life cycle or the transition from sexual immaturity to maturity.

**De novo telomere addition**

The presence of DNA double-strand breaks (DSBs) is mostly known to be associated with the induction of repair machinery [Non-Homologous End Joining (NHEJ) or Homologous Recombination (HR)]. The relatively large number of DSBs associated with chromosome fragmentation during development in ciliates generates a large number of chromosomes whose broken ends are “healed” through de novo telomere addition. As in other eukaryotes, ciliate telomeres consist of tandem repeats at the 5′ and 3′ ends of their chromosomes, such as 5′-GGGTTTTT-3′ (G₄T₄) in *Oxytricha* and *Euplotes* [168], 5′-GGGGTT-3′ (G₁T₂) in *Tetrahymena* [22] and 5′-GGGGTT-3′ or 5′-GGGTTT-3′ in *Paramecium* [169, 170].

Although chromosome fragmentation is a very reproducible and relatively precise event, de novo telomere addition does not typically occur at precise nucleotide position (the exception being *E. crassus* [158]), generating microheterogeneity among the amplified chromosome copies in the developing macronucleus. In *Oxytricha* and *Tetrahymena*, telomere addition sites have been found to be clustered within regions ≤30 bp, [165, 171–173], whereas this is often ~1 to 2 Kbp in *Paramecium* [169, 170]. However, in *E. crassus*, there is no heterogeneity and telomeres are added at the same nucleotide positions in all macronuclear copies [158]. Additional heterogeneity can arise from the use of alternative chromosome fragmentation sites in *Paramecium* and *Oxytricha*. In *Paramecium*, the ends of some MAC chromosomes can be generated at alternative telomere addition sites separated by 2–13 Kbp. Each of the regions shows heterogeneity in the telomere’s positions [170, 174, 175]. While most of the chromosome fragmentation in *Oxytricha* results in gene-sized chromosomes, the use of alternative fragmentation sites (or failure to fragment) can result in macronuclear chromosomes encoding additional ORFs [154, 172, 173].

The exact mechanism of de novo telomere addition remains poorly understood. Data performed on *Tetrahymena* suggested the involvement of the telomere end binding homologue Pot2p in de novo telomere addition that exclusively localizes to CBSs during chromosome fragmentation [176]. Recent work from *Stylonychia* has shown that microinjection of RNA templates carrying variable telomeric repeats into the developing macronucleus leads to modified telomeres in vegetative cells suggesting that de novo telomere addition depends on a telomere-containing transcript derived from the parental macronucleus [177]. However, to understand this process in more detail, further work needs to be done.

**Conclusions**

Ciliates are a diverse group of organisms that have deeply contributed to our recent knowledge about the regulatory role of epigenetics in development. Identification of sRNA pathways as well as histone modifications that mediate DNA elimination is providing a greater understanding of the genome reorganization process in ciliates while shedding new insight into the evolution of epigenetic processes across eukaryotes. While we have a basic understanding of the overall genome reorganization process, numerous outstanding questions remain open. How are IES regions identified and preferentially transcribed to produce sRNAs in the meiotic MIC? What, if any, role is there for retaining IESs in the MIC? How has this process evolved across the ciliate phylogeny? Further comparative analyses of somatic and germline genomes, and the associated DNA elimination process, will be instrumental in answering these questions and will ultimately shed light on the variety of RNA-mediated epigenetic pathways and the dynamic regulation of genome function and structure.

**Acknowledgements** This work was supported by the funding 500-3-11-110. We would like to thank Fabian Wieland for help with Fig. 2 as well as the reviewers for valuable comments.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, and indicate if changes were made.

To view a copy of this licence, visit [http://creativecommons.org/licenses/by/4.0/](http://creativecommons.org/licenses/by/4.0/).
References

1. Streit A, Wang J, Kang Y, Davis RE (2016) Gene silencing and sex determination by programmed DNA elimination in parasitic nematodes. Curr Opin Microbiol 32:120–127
2. Kloc M, Zagrodzinska B (2001) Chromatin elimination—an oddity or a common mechanism in differentiation and development? Differentiation 68:84–91. https://doi.org/10.1046/j.1432-0436.2001.680202.x
3. Wang J, Davis RE (2014) Programmed DNA elimination in multicellular organisms. Curr Opin Genet Dev 27:26–34
4. Boveri T (1887) Ueber Differenzierung der Zellkerne wahrend der Furchung des Eies von Ascaris megaloecephala. The first discovery and description of a DNA elimination, chromatin diminution, in the parasitic nematode Parascaris. Anat Anz 2:688–693
5. Bassing CH, Swat W, Alt FW (2002) The mechanism and regulation of chromosomal V(D)J recombination. Cell 109:45–55
6. Parfrey LW, Lahr DJG, Knoll AH, Katz LA (2011) Estimating the timing of early eukaryotic diversification with multigene molecular clocks. Proc Natl Acad Sci 108:13624–13629. https://doi.org/10.1073/pnas.110633108
7. Chalker DL, Yao M-C (2011) DNA elimination in ciliates: transposon domestication and genome surveillance. Annu Rev Genet 45:227–246. https://doi.org/10.1146/annurev-genet-110410-132342
8. Bracht JR, Fang W, Goldman AD et al (2013) Genomes on the edge: programmed genome instability in ciliates. Cell 152:406–416
9. Zufall RA, Robinson T, Katz LA (2005) Evolution of developmentally regulated genome rearrangements in eukaryotes. J Exp Zool B Mol Dev Evol 304:448–455
10. Grant JR, Katz LA (2014) Building a phylogenomic pipeline for the eukaryotic tree of life—addressing deep phylogenies with genome-scale data. PLoS Curr 6. Doi: 10.1371/currents.tol.c24b6054aebf3602748ac042ccc8f2e9
11. Yan Y, Rogers AJ, Gao F, Katz LA (2017) Unusual features of nuclear division. Mol Phylogenet Evol 70:240–243. https://doi.org/10.1016/j.ympev.2013.10.001
12. Foisner W, Chao A, Katz LA (2008) Diversity and geographic distribution of ciliates (Protista: Ciliophora). Biodivers Conserv. 17:345–363
13. Simon M, Plattner H (2014) Unicellular eukaryotes as models for investigating differentiation and development? Differentiation 68:84–91. https://doi.org/10.1046/j.1432-0436.2001.680202.x
14. Lozupone CA, Knight RD, Landweber LF (2001) The molecular basis of nuclear genetic code change in ciliates. Curr Biol 11:65–74. https://doi.org/10.1016/S0960-9822(01)00028-8
15. Swart EC, Serra V, Petroni G, Nowacki M (2016) Genetic codes with no dedicated stop codon: context-dependent translation termination. Cell 166:691–702. https://doi.org/10.1016/j.cell.2016.06.020
16. Lipton IH (2013) Nuclear codes with no dedicated stop codons. Cell 155:1–8. https://doi.org/10.1016/j.cell.2013.12.017
17. Doerder FP, De Bault LE (1975) Cytofluorometric analysis of nuclear DNA during meiosis, fertilization and macronuclear development in the ciliate Stichodactyla. Curr Biol 70:211–223. https://doi.org/10.1016/j.cub.2014.07.034
18. Yan Y, Rogers AJ, Gao F, Katz LA (2017) Unusual features of nuclear DNA recombination in the ciliate Paramaecium. J Cell Sci 129:406–416. https://doi.org/10.1242/jcs.182184
19. Swart EC, Serra V, Petroni G, Nowacki M (2016) Genetic codes with no dedicated stop codon: context-dependent translation termination. Cell 166:691–702. https://doi.org/10.1016/j.cell.2016.06.020
20. Heaphy SM, Mariotti M, Gladyshev VN et al (2016) Novel ciliate genetic code variants including the reassignment of all three stop codons to sense codons in condylostoma magnum. Mol Biol Evol 33:2885–2889. https://doi.org/10.1093/molbev/msw166
21. Greider CW, Blackburn EH (1985) Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. Cell 43:405–413. https://doi.org/10.1016/0009-8674(85)90170-0
22. Blackburn EH, Gall JG (1978) A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. J Mol Biol 120:33–53. https://doi.org/10.1016/0092-8674(78)90294-2
23. 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. https://doi.org/10.1016/S0092-8674(02)00909-1
24. Taverna SD, Coyne RS, Allis CD (2002) Methylation of histone H3 at lysine 9 targets programmed DNA elimination in Tetrahymena. Cell 110:701–711. https://doi.org/10.1016/S0092-8674(02)00941-8
25. Allen SE, Hug I, Pabian S et al (2017) Circular Concatemers of Ultra-Short DNA Segments Produce Regulatory RNAs. Cell 168:990–999. https://doi.org/10.1016/j.cell.2017.02.020
26. Prescott DM (1994) The DNA of ciliated protozoa. Microbiol Rev 58:233–267. https://doi.org/10.1128/972.00095-2
27. Barid SE, Klobucher LA (1988) Genetic characterization and use of a restriction fragment length variant in the hypotrichous ciliate euplotes crassus. J Protozool 35:459–465. https://doi.org/10.1111/j.1550-7408.1988.tb04130.x
28. Godiska R, Aufderheide KD, Gilley D et al (1987) Transformation of paramecium by microinjection of a cloned serotype gene. Proc Natl Acad Sci USA 84:7590–7594. https://doi.org/10.1073/pnas.84.21.7590
29. Martindale D, Martindale H, BP, (1986) Nucleic acids research. Nucleic Acids Res 14:1341–1354. https://doi.org/10.1093/nar/gkh411
30. Hamilton EP, Kapusta A, Huvos PE et al (2016) Structure of the germline genome of Tetrahymena thermophila and relationship to the massively rearranged somatic genome. Elife 5:e19090. https://doi.org/10.7554/eLife.19090.001
31. Chen X, Bracht JR, Goldman AD et al (2014) The architecture of a scrambled genome reveals massive levels of genomic rearrangement during development. Cell 158:1187–1198. https://doi.org/10.1016/j.cell.2014.07.034
32. Yan Y, Rogers AJ, Gao F, Katz LA (2017) Unusual features of non-dividing somatic macronuclei in the ciliate class Karyorelictea. Eur J Protop 61:399–408
33. Wancura MM, Yan Y, Katz LA, Maurer-Alcalá XX (2018) Nuclear features of the heterotrich ciliate blepharisma americanum: genomic amplification, life cycle, and nuclear inclusion. J Eukaryot Microbiol 65:4–11. https://doi.org/10.1111/jeu.12422
34. Juraneck SA, Lippis HJ (2007) New Insights into the Macronuclear Development in Ciliates. Int Rev Cytol 252:219–251
35. Doerder FP, De Bault LE (1975) Cytofluorometric analysis of nuclear DNA during meiosis, fertilization and macronuclear development. Curr Opin Genet Dev 70:211–223. https://doi.org/10.1016/j.cub.2014.07.034
36. Seyfert HM, Cleffmann G (1982) Mean macronuclear DNA content in ciliates. Curr Opin Genet Dev 70:211–223. https://doi.org/10.1016/j.cub.2014.07.034
37. Seyfert HM, Cleffmann G (1982) Mean macronuclear DNA content in ciliates. Curr Opin Genet Dev 70:211–223. https://doi.org/10.1016/j.cub.2014.07.034
37. Doerder FP, Deak JC, Lief JH (1992) Rate of phenotypic assortment in Tetrahymena thermophila. Dev Genet 13:126–132. https://doi.org/10.1002/dvg.1201032006

38. Eisen JA, Coyne RS, Wu M et al (2006) Macronuclear genome sequence of the ciliate Tetrahymena thermophila, a model eukaryote. PLoS Biol 4:e286. https://doi.org/10.1371/journal.pbiol.0040286

39. Aury JM, Jaillon O, Duret L et al (2006) Global trends of whole-genome duplications revealed by the ciliate Paramecium tetraurelia. Nature 444:171–178. https://doi.org/10.1038/nature05230

40. Duret L, Cohen J, Jubin C et al (2008) Analysis of sequence variability in the macronuclear DNA of Paramecium tetraurelia: A somatic view of the germline. Genome Res 18:585–596. https://doi.org/10.1101/gr.074534.107

41. Nowacki M, Haye JE, Fang W et al (2010) RNA-mediated epigenetic regulation of DNA copy number. Proc Natl Acad Sci 12:367–389. https://doi.org/10.1073/pnas.1012236107

42. Xu K, Doak TG, Lipps HJ et al (2012) Copy number variations of 11 macronuclear chromosomes and their gene expression in Oxytricha trifallax. Gene 505:75–80. https://doi.org/10.1016/j.gene.2012.05.045

43. Swart EC, Bracht JR, Magrini V et al (2013) The Oxytricha trifallax Macronuclear Genome: A Complex Eukaryotic Genome with 16,000 Tiny Chromosomes. PLoS Biol 11:e1001473. https://doi.org/10.1371/journal.pbio.1001473

44. Sonneborn TM (1974) Paramecium aurelia. In: Handbook of Eukaryotic Microbiology. pp 442–452.

45. Jacobs ME, Klobutcher LA (1996) The long and the short of developmental dna deletion in Euplotes crassus. In: Journal of Eukaryotic Microbiology. pp 442–452.

46. Klobutcher LA, Herrick G (1995) Consensus inverted terminal repeat sequence of Paramecium IESs: Resemblance to termini of Tc1-related and Euplotes Tec transposons. Nucleic Acids Res 23:2006–2013. https://doi.org/10.1093/nar/23.11.2006

47. Arnaiz O, Mathy N, Baudry C et al (2012) The Paramecium Germline Genome Provides a Niche for Intragenic Parasitic DNA: Evolutionary Dynamics of Internal Eliminated Sequences. PLoS Genet 8:e1002984. https://doi.org/10.1371/journal.pgen.1002984

48. Jaraczewski JW, John CL (1993) Elimination of Tec elements involves a novel excision process. Genes Dev 7:95–105. https://doi.org/10.1101/gad.7.1.95

49. Guérin F, Arnaiz O, Boggetto N et al (2017) Flow cytometry analysis of Programmed DNA Elimination Sites in Tetrahymena thermophila. G3|Genes|Genomes|Genetics 1:515–522. Doi:10.1534/g3.111.000927

50. Klobutcher LA, Herrick G (1997) Developmental Genome Reorganization in Ciliated Protozoa: The Transposon Link. pp 1–62.

51. Guérin F, Arnaiz O, Boggetto N et al (2017) Flow cytometry sorting of nuclei enables the first global characterization of Paramecium germline DNA and transposable elements. BMC Genomics 18:327. https://doi.org/10.1186/s12864-017-3713-7

52. Maurer-Alcalá XX, Yan Y, Pilling OA et al (2018) Twisted Tales: Insights into Genome Diversity of Ciliates Using Single-Cell ’Omics. Genome Biol Evol 10:1927–1939. https://doi.org/10.1093/gbe/eyy133

53. Maurer-Alcalá XX, Knight R, Katz LA (2018) Exploration of the germline genome of the ciliate Chilodonella uncinata within single-cell omics (Transcriptomics and genomics). MBio 9:e01836–e1917. https://doi.org/10.1128/mBio.01836-17

54. Callahan RC, Shalke G, Gorovsky MA (1984) Developmental rearrangements associated with a single type of expressed α-tubulin gene in tetrahymena. Cell 36:441–445. https://doi.org/10.1016/0092-8674(84)90237-X

55. Yao MC, Gorovsky MA (1974) Comparison of the sequences of macro- and micronuclear DNA of Tetrahymena pyriformis. Chromosoma 48:1–18. https://doi.org/10.1007/BF00284863

56. Yao MC, Choi J, Yokoyama S et al (1984) DNA elimination in tetrahymena: A developmental process involving extensive breakage and rejoining of DNA at defined sites. Cell 36:433–440. https://doi.org/10.1016/0092-8674(84)90236-8

57. Gratias A, Bettermier M (2003) Processing of Double-Strand Breaks Is Involved in the Precise Excision of Paramecium Internal Eliminated Sequences. Mol Cell Biol 23:7152–7162. https://doi.org/10.1128/mcb.23.20.7152-7162.2003

58. Eder C, Maercker C, Meyer J, Lipps HJ (1993) The processing of macronuclear DNA sequences during macronuclear development of the hypotrichous ciliate Stylonychia lemmiae. Int J Dev Biol 37:473–477. https://doi.org/10.1515/ijdb.8292542

59. Prescott DM, DuBois ML (1996) Internal eliminated segments (IESs) of oxytrichidae. In: Journal of Eukaryotic Microbiology. pp 432–441.

60. Heinonen TYK, Pearlman RE (1994) A germ line-specific sequence element in an intron in Tetrahymena thermophila. J Biol Chem 269:17428–17433

61. Chilcoat ND, Turkewitz AP (1997) In vivo analysis of the major exocytosis-sensitive phosphoprotein in Tetrahymena. J Cell Biol 139:1197–1207. https://doi.org/10.1083/jcb.139.5.1197

62. Fass JN, Joshi NA, Couvillion MT, et al (2011) Genome-Scale Analysis of Programmed DNA Elimination Sites in Tetrahymena thermophila. G3Genes|Genomes|Genetics 1:515–522. doi:10.1534/g3.111.000927

63. Le Mouël A, Butler A, Caron F, Meyer E (2003) Developmentally regulated chromosome fragmentation linked to precise elimination of repeated sequences in Paramecia. Eukaryot Cell 2:1076–1090. https://doi.org/10.1128/EC.2.5.1076-1090.2003

64. Catania F, McGrath CL, Doak TG, Lynch M (2013) Spliced DNA sequences in the Paramecium germline: Their properties and evolutionary potential. Genome Biol Evol 5:1200–1211. https://doi.org/10.1093/gbe/evt087

65. Mayer KM, Mikami K, Forney JD (1998) A mutation in Paramecium tetraurelia reveals functional and structural features of developmentally excised DNA elements. Genetics 148:139–149

66. Mayer KM, Forney JD (1999) A mutation in the flanking 5’-TA-3’ dinucleotide prevents excision of an internal eliminated sequence from the Paramecium tetraurelia genome. Genetics 151:597–694

67. Gratias A, Lepère G, Garnier O et al (2008) Developmentally programmed DNA splicing in Paramecium reveals short-distance crosstalk between DNA cleavage sites. Nucleic Acids Res 36:3244–3251. https://doi.org/10.1093/nar/gkn154

68. Ku M, Mayer K, Forney JD (2000) Developmentally regulated excision of a 28-base-pair sequence from the Paramecium genome requires flanking DNA. Mol Cell Biol 20:8390–8396. https://doi.org/10.1128/MCB.20.22.8390-8396.2000

69. Bhullar S, Wilkes CD, Arnaiz O et al (2018) A mating-type mutagenesis screen identifies a zinc-finger protein required for specific DNA excision events in Paramecium. Nucleic Acids Res 46:9550–9562. https://doi.org/10.1093/nar/gky772

70. Chalker DL, La Terza A, Wilson A et al (1999) Flanking regulatory sequences of the Tetrahymena R deletion element determine the boundaries of DNA rearrangement. Mol Cell Biol 19:5631–5641

71. Godiska R, James C, Yao MC (1993) A distant 10-bp sequence specifies the boundaries of a programmed DNA deletion in Tetrahymena. Genes Dev 7:2357–2365. https://doi.org/10.1101/gad.7.12a.2357

72. Lin C-YG, Chao J-L, Tsai H-K et al (2019) Setting boundaries for genome-wide heterochromatic DNA deletions through flanking inverted repeats in Tetrahymena thermophila. Nucleic Acids Res 47:5181–5192. https://doi.org/10.1093/nar/gkz209

73. Jaspans VN, Taye ME, Carle CM et al (2019) Boundaries of eliminated heterochromatin of Tetrahymena are positioned by
the DNA-binding protein Ltl1. Nucleic Acids Res 14:7348–7362. https://doi.org/10.1093/nar/gzk504
74. Carle CM, Zaher HS, Chalker DL (2016) A Parallel G Quadruplex-Binding Protein Regulates the Boundaries of DNA Elimination Events of Tetrahymena thermophila. PLoS Genet 12:e1005842. https://doi.org/10.1371/journal.pgen.1005842
75. Swart EC, Wilkes CD, Sandoval PY et al (2014) Genome-wide analysis of genetic and epigenetic control of programmed DNA deletion. Nucleic Acids Res 42:8970–8983. https://doi.org/10.1093/nar/gku619
76. Sandoval P, Swart E, Arambasic M, Nowacki M (2014) Functional Diversification of Dicer-like Proteins and Small RNAs Required for Genome Sculpting. Dev Cell 28:174–188. https://doi.org/10.1016/j.devcel.2013.12.010
77. Kapusta A, Matsuda S, Marmignon A et al (2011) Highly precise and developmentally programmed genome assembly in paramecium requires ligase IV-dependent end joining. PLoS Genet 7:e1002049. https://doi.org/10.1371/journal.pgen.1002049
78. Lin I-T, Chao J-L, Yao M-C (2012) An essential role for the DNA-binding protein Ltl1. Nucleic Acids Res 14:7348–7362. https://doi.org/10.1093/nar/gku619
79. Marmignon A, Bischerour J, Silve A et al (2014) Ku-Mediated Breakage-repair protein Ku80 in programmed DNA rearrangement. PLoS Genet 10:e1004552. https://doi.org/10.1371/journal.pgen.1004552
80. Mochizuki K, Gorovsky MA (2004) RNA polymerase II localizes to the DNA elimination site in Tetrahymena. Proc Natl Acad Sci U S A 111:1233–1238. https://doi.org/10.1073/pnas.0507682102
81. Hoffmann DC, Prescott DM (1997) Phylogenetic relationships among hypotrichous ciliates determined with the macronuclear gene encoding the large, catalytic subunit of DNA polymerase α. J Mol Evol 45:301–310. https://doi.org/10.1007/PL00006234
82. Hoffmann D (1996) The germline gene encoding DNA polymerase alpha in the hypotrichous ciliate Oxytricha nova is extremely scrambled. Nucleic Acids Res 24:301–310. https://doi.org/10.1093/nar/24.17.3337
83. Malone CD, Anderson AM, Motl JA et al (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. https://doi.org/10.1128/MCB.25.20.9151–9164.2005
84. 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. https://doi.org/10.1101/gad.1265105
85. Chang W-J, Bryson PD, Liang H et al (2005) The evolutionarily ancient origin of a complex scrambled gene. Proc Natl Acad Sci USA 102:15149–15154. https://doi.org/10.1073/pnas.0507682102
86. Hoffman DC, Prescott DM (1997) Phylogenetic relationships among hypotrichous ciliates determined with the macronuclear gene encoding the large, catalytic subunit of DNA polymerase α. J Mol Evol 45:301–310. https://doi.org/10.1007/PL00006234
87. Hoffman D (1996) The germline gene encoding DNA polymerase alpha in the hypotrichous ciliate Oxytricha nova is extremely scrambled. Nucleic Acids Res 24:301–310. https://doi.org/10.1093/nar/24.17.3337
88. Landweber LF, Kuo TC, Curtis EA (2000) Evolution and assembly of an extremely scrambled gene. Proc Natl Acad Sci USA 97:3298–3303. https://doi.org/10.1073/pnas.97.7.3298
89. Nowacki M, Vijayan V, Zhou Y et al (2008) RNA-mediated epigenetic programming of a genome-rearrangement pathway. Nature 451:153–158. https://doi.org/10.1038/nature06452
106. Lindblad KA, Bracht JR, Williams AE, Landweber LF (2017) Thousands of RNA-cached copies of whole chromosomes are present in the ciliate Oxytricha during development. RNA 23:1200–1208. https://doi.org/10.1261/rna.058511.116

107. Khurana JS, Wang X, Chen X et al (2014) Transcription-independent functions of an RNA polymerase II subunit, Rpb2, during genome rearrangement in the ciliate, Oxytricha trifallax. Genetics 197:839–849. https://doi.org/10.1534/genetics.114.163279

108. Kari L, Landweber LF (1999) Computational Power of Gene Rearrangement. Proceedings of DNA Based Computers V, E. Winfree, D. Gifford eds., MIT, Boston 0000:207–216.

109. Ehrenfeucht A, Harju T, Rozenberg G (2002) Gene assembly through cyclic graph decomposition. Theor Comput Sci 281:325–349. https://doi.org/10.1016/S0304-3975(02)00019-1

110. Liu Y, Mochizuki K, Gorovsky MA (2004) Histone H3 lysine 9 methylation is required for DNA elimination in developing macronuclei in Tetrhydema. Proc Natl Acad Sci U S A 101:1679–1684. https://doi.org/10.1073/pnas.0305421101

111. Liu Y, Taverna SD, Muratore TL et al (2007) RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in Tetrhydema. Genes Dev 21:1530–1545. https://doi.org/10.1101/gad.1544207

112. Frapporti A, Miró Pina C, Arnaiz O et al (2019) The Polycytoplasm protein Ez11 mediates H3K9 and H3K27 methylation to repress transposable elements in Paramecium. Nat Commun 10:2710. https://doi.org/10.1038/s41467-019-10648-5

113. Papazyan R, Voronina E, Chapman JR et al (2014) Methylation of histone H3K23 blocks DNA damage in pericentric heterochromatin during meiosis. Elife 3:e02996. https://doi.org/10.7554/elife.02996

114. Suhren JH, Noto T, Kataoka K et al (2017) Negative Regulators of an RNAi-Heterochromatin Positive Feedback Loop Safeguard Somatic Genome Integrity in Tetrhydema. Cell Rep 18:2494–2507. https://doi.org/10.1016/j.celrep.2017.02.024

115. Kataoka K, Mochizuki K (2015) Phosphorylation of an HP1-like Protein Regulates Heterochromatin Body Assembly for DNA Elimination. Dev Cell 35:775–788. https://doi.org/10.1016/j.devcel.2015.11.017

116. Cheng CY, Vogt A, Mochizuki K, Yao MC (2010) A Domesticated piggyBac Transposable Plays Key Roles in Heterochromatin Dynamics and DNA Cleavage during Programmed DNA Deletion in Tetrhydema thermophila. Mol Biol Cell 21:1753–1762. https://doi.org/10.1091/mbc.E09-12-1079

117. Zhao X, Xiong J, Mao F et al (2019) RNAi-dependent polycytoplasm repression controls transposable elements in Tetrhydema. Genes Dev 33:348–364. https://doi.org/10.1101/gad.320796.118

118. Taverna SD, Ueberheide BM, Liu Y et al (2007) Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. Proc Natl Acad Sci U S A 104:2086–2091. https://doi.org/10.1073/pnas.0610993104

119. Garcia BA, Hake SB, Diaz RL et al (2007) Organismal differences in post-translational modifications in histones H3 and H4. J Biol Chem 282:7641–7655. https://doi.org/10.1074/jbc.M607900200

120. Lhuillier-Akakpo M, Frapport A, Denby Wilkes C et al (2014) Local Effect of Enhancer of Zeste-Like Reveals Cooperation of Epigenetic and cis-Acting Determinants for Zygotic Genome Rearrangements. PLoS Genet 10:e1004665. https://doi.org/10.1371/journal.pgen.1004665

121. Ignarski M, Singh A, Swart EC et al (2014) Paramecium tetrarelia chromatin assembly factor-1-like protein PcCAF-1 is involved in RNA-mediated control of DNA elimination. Nucleic Acids Res 42:11952–11964. https://doi.org/10.1093/nar/gku874

122. Ammermann D (1971) Morphology and development of the macronuclei of the ciliates Stylonychia mytilus and Euplotes aediculatus. Chromosoma 33:209–238. https://doi.org/10.1007/BF00285634

123. Forcob S, Bulic A, Jonsson F et al (2014) Differential expression of histone H3 genes and selective association of the variant H3.7 with a specific sequence class in Stylonychia macronuclear development. Epigenetics and Chromatin. Doi: 10.1186/1756-8935-7-4

124. Postberg J, Heyse K, Cremer M et al (2008) Spatial and temporal plasticity of chromatin during programmed DNA-reorganization in Stylonychia macronuclear development. Epigenetics Chromatin 1:3. https://doi.org/10.1186/1756-8935-1-3

125. Jurane SA, Rupprecht S, Postberg J, Lipps HJ (2005) snRNA and heterochromatin formation are involved in DNA excision during macronuclear development in stichotrichous ciliates. Eukaryot Cell 4:1934–1941. https://doi.org/10.1128/EC.4.11.1934-1941.2005

126. Meyer GF, Lipps HJ (1980) Chromatin elimination in the hypotrichous ciliate Stylonychia mytilus. Chromosoma 77:285–297. https://doi.org/10.1007/BF00286054

127. Bracht JR, Perlman DH, Landweber LF (2012) Cytosine methylation and hydroxymethylation mark DNA for elimination in Oxytricha trifallax. Genome Biol 13:R99. https://doi.org/10.1186/gb-2012-13-10-99

128. Jurane S, Wieden H, Lipps HJ (2003) De novo cytosine methylation in the differentiating macronucleus of the stichotrichous ciliate Stylonychia lemnacea. Nucleic Acids Res 31:1387–1391

129. Bulic A, Postberg J, Fischer A et al (2013) A permissive chromatin structure is adopted prior to site-specific DNA demethylation of developmentally expressed genes involved in macronuclear differentiation. Epigenetics and Chromatin 6:5. https://doi.org/10.1186/1756-8935-6-5

130. Singh A, Vancara A, Woycicki RK et al (2018) Determination of the presence of 5-methylcytosine in Paramecium tetraurelia. PLoS ONE 13:e0206667. https://doi.org/10.1371/journal.pone.0206667

131. Harrison GS, FindlyKarrer RCKM (1986) Site-specific methylation of adenine in the nuclear genome of a eucaryote. Tetrahymena thermophila Mol Cell Biol 6:2364–2370. https://doi.org/10.1128/MCB.6.7.2364

132. Cummings DJ, Tait A, Goddard LF (1974) Methylated bases in DNA from Paramecium aurelia. BBA Sect Nucleic Acids Protein Synth 374:1–11. https://doi.org/10.1016/0005-2787(74)90194-4

133. Wang Y, Chen X, Sheng Y et al (2017) N6-adenine DNA methylation is associated with the linker DNA of H2A.Z-containing well-positioned nucleosomes in Pol II-transcribed genes in Tetrahymena. Nucleic Acids Res 45:11594–11606. https://doi.org/10.1093/nar/gkx883

134. Gorovsky MA, Hattman S, Pleger GL (1973) [6N]methyl adenine in the nuclear DNA of a eucaryote. Tetrahymena Pyriformis J Cell Biol 56:697–701. https://doi.org/10.1083/jcb.56.3.697

135. Beh LY, Debelouchina GT, Clay DM et al (2019) Identification of a DNA N6-adenine methyltransferase complex and its impact on chromatin organization. Cell 177:1781–1796. https://doi.org/10.1016/j.cell.2019.04.028

136. Luo GZ, Hao Z, Luo L et al (2018) N 6 -methyldeoxyadenosine directs nucleosome positioning in Tetrahymena DNA, Genome Biol 19:200. https://doi.org/10.1186/s13059-018-1573-3

137. Wang Y, Sheng Y, Liu Y et al (2017) N6-methyladenine DNA modification in the unicellular eukaryotic organism Tetrahymena thermophila. Eur J Proteol 58:94–102. https://doi.org/10.1016/j.ejop.2016.12.003

138. Karrer KM (2002) Methylation of adenine in the nuclear DNA of Tetrahymena is internucleosomal and independent of histone
139. Wang Y, Sheng Y, Liu Y et al (2019) A distinct class of eukaryotic MT-A70 methyltransferases maintain symmetric DNA N6-adenine methylation at the ApT dinucleotides as an epi-genetic mark associated with transcription. Nucleic Acids Res 47:11771–11789. https://doi.org/10.1093/nar/gkz1053

140. Pancer Z, Amemiya CT, Ehrhardt GRA et al (2004) Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. Nature 430:174–180. https://doi.org/10.1038/nature02740

141. Flajnik MF (2004) Another manifestation of GOD. Nature 430:157–158

142. Baudry C, Malinsky S, Restituito M et al (2009) PiggyMac, a domesticated piggybac transposase together carry out programmed DNA elimination in tetrahymena. PLoS Genet 9:e1004032. https://doi.org/10.1371/journal.pgen.1004032

143. Vogt A, Mochizuki K (2013) A domesticated piggyBac transposase interacts with heterochromatin and catalyzes reproducible DNA elimination in tetrahymena. PLoS ONE 8:e75337. https://doi.org/10.1371/journal.pone.0075337

144. Mitra R, Fain-Thornton J, Craig NL (2008) PiggyBac can bypass DNA synthesis during cut and paste transposition. EMBO J 27:1097–1109. https://doi.org/10.1038/emboj.2008.41

145. Bischlerour J, Bhullar S, Wilkes CD et al (2018) Six domesticated piggybac transposases together carry out programmed DNA elimination in paramecium. Elife 7:e37927. https://doi.org/10.7554/eLife.37927

146. Shieh AWY, Chalker DL (2013) LIAS Is Required for nuclear reorganization and programmed DNA rearrangements occurring during tetrahymena macronuclear differentiation. PLoS ONE 8:e75337. https://doi.org/10.1371/journal.pone.0075337

147. Cheng CY, Young JM, Lin CYG et al (2016) The piggyBac transposon-derived genes TPB1 and TPB6 mediate essential transposon-like excision during the developmental rearrangement of key genes in Tetrahymena thermophila. Genes Dev 30:2724–2736. https://doi.org/10.1101/gad.290460.116

148. Feng L, Wang G, Hamilton EP et al (2017) A germline-limited piggyBac transposase gene is required for precise excision in Tetrahymena genome rearrangement. Nucleic Acids Res 45:9481–9502. https://doi.org/10.1093/nar/gkx652

149. Nowacki M, Higgins BP, Maqilman GM et al (2009) A functional role for transposases in a large eukaryotic genome. Science (80-) 324:935–938. https://doi.org/10.1126/science.1170023

150. Hamilton E, Bruns P, Lin C et al (2005) Genome-wide characterization of Tetrahymena thermophila telomere addition sites in Euplotes crassus. Nucleic Acids Res 32:4230–4240. https://doi.org/10.1093/nar/gsi186

151. Chen X, Landweber LF (2016) Phylogenomic analysis reveals genome-wide purifying selection on TBE transposons in the ciliated protozoa and a common “D35E” motif. Proc Natl Acad Sci USA 113:6957–6962. https://doi.org/10.1073/pnas.1603684113

152. Hunter DJ, Williams K, Cartinhour S, Herrick G (1989) Precise nuclear genome rearrangements in the ciliate Paramecium tetraurelia. Mol Cell 8:251–258. https://doi.org/10.1016/0083-4421(89)90142-2

153. Aeschlimann SH, Jönsson F, Postberg J et al (2014) The draft genome of the sea lamprey. Nature 430:174–180. https://doi.org/10.1038/nature13887

154. Seegmiller A, Williams KR, Herrick G (1997) Two two-gene candidate fragmentation sequence in Euplotes crassus: a transposonlike element is removed during macronuclear development. Mol Cell Biol 9:3793–3807. https://doi.org/10.1128/MCB.9.9.3793

155. Conover RK, Brunk CF (1986) Macronuclear DNA molecules of Tetrahymena thermophila. Mol Cell Biol 6:900–905. https://doi.org/10.1128/MCB.6.3.900

156. Phan HL, Forney J, Blackburn EH (1989) Analysis of paramecium macronuclear DNA using pulsed field gel electrophoresis. J Protopozool 36:402–408. https://doi.org/10.1111/j.1550-7408.1989.tb05535.x

157. Yao MC, Yao CH, Monks B (1990) The controlling sequence for site-specific chromosome breakage in tetrahymena. Cell 63:763–772. https://doi.org/10.1016/0092-8674(90)90142-2

158. Baird SE, Klobutcher LA (1989) Characterization of chromosome fragmentation in two protozoans and identification of a candidate fragmentation sequence in Euplotes crassus. Genes Dev 3:585–597. https://doi.org/10.1101/gad.3.5.585

159. Klobutcher LA (1999) Characterization of in vivo developmental chromosome fragmentation intermediates in E. crassus. Mol Cell 4:695–704. https://doi.org/10.1016/S1097-2765(00)80380-9

160. Lin CYG, Lin IT, Yao MC (2016) Programmed Minichromosome elimination as a mechanism for somatic genome reduction in Tetrahymena thermophila. Nucleic Acids Res 28:895–900

161. Hamilton EP, Williamson S, Dunn S et al (2006) The highly conserved family of Tetrahymena thermophila chromosome breakage elements contains an invariant 10-base-pair core. Eukaryot Cell 5:771–780. https://doi.org/10.1128/EC.5.4.771-780.2006

162. Fan Q, Yao MC (2000) A long stringent sequence signal for programmed chromosome breakage in Tetrahymena thermophila. Nucleic Acids Res 28:895–900

163. Hamilton EP, Williamson S, Dunn S et al (2006) The highly conserved family of Tetrahymena thermophila chromosome breakage elements contains an invariant 10-base-pair core. Eukaryot Cell 5:771–780. https://doi.org/10.1128/EC.5.4.771-780.2006

164. Lin CYG, Lin IT, Yao MC (2016) Programmed Minichromosome elimination as a mechanism for somatic genome reduction in Tetrahymena thermophila. PLoS Genet 12:e1006403. https://doi.org/10.1371/journal.pgen.1006403

165. Fan Q, Yao M (1996) New telomere formation coupled with site-specific chromosome breakage in Tetrahymena thermophila. Mol Cell Biol 16:1267–1274. https://doi.org/10.1128/MCB.16.3.1267

166. Cassidy-Hanley D, Bisharyan Y, Fridman V et al (2005) Genome-wide characterization of Tetrahymena thermophila chromosome breakage sites. II Phys Genet Mapping Genet 170:1623–1631. https://doi.org/10.1534/genetics.104.031435

167. Hamilton E, Bruns P, Lin C et al (2005) Genome-wide characterization of Tetrahymena thermophila chromosome breakage sites. I. Cloning and identification of functional sites. Genetics 170:1611–1621. https://doi.org/10.1034/j.1553-4408.2011.14001

168. Klobutcher LA, Swanton MT, Donini P, Prescott DM (1981) All telomere addition in wild-type and mutant paramecia. Mol Cell Biol 1:1047–1058. https://doi.org/10.1016/0092-8674(81)90142-2

169. Hamilton E, Bruns P, Lin C et al (2005) Genome-wide characterization of Tetrahymena thermophila chromosome breakage sites. I. Cloning and identification of functional sites. Genetics 170:1611–1621. https://doi.org/10.1034/j.1553-4408.2011.14001

170. Baird SE, Fino GM, Tausta SL, Klobutcher LA (1989) Macronuclear genome organization in Euplotes crassus: a transposonlike element is removed during macronuclear development. Mol Cell Biol 9:3793–3807. https://doi.org/10.1128/MCB.9.9.3793

171. Seegmiller A, Williams KR, Herrick G (1997) Two two-gene macronuclear chromosomes of the hypotrichous ciliates Oxytricha fallax and O. trifallax generated by alternative processing of the 81 locus. Dev Genet 20:348–357. https://doi.org/10.1002/(SICI)1520-6408(1997)20:4%3c348:AID-DVG6%3e3.0.CO;2-7
Programmed genome rearrangements in ciliates

173. Williams KR, Doak TG, Herrick G (2002) Telomere formation on macronuclear chromosomes of Oxytricha trifallax and O. fallax: Alternatively processed regions have multiple telomere addition sites. BMC Genet. 3:16

174. Amar L (1994) Chromosome end formation and internal sequence elimination as alternative genomic rearrangements in the ciliate paramecium. J Mol Biol 236:421–426

175. Caron F (1992) A high degree of macronuclear chromosome polymorphism is generated by variable DNA rearrangements in Paramecium primaurelia during macronuclear differentiation. J Mol Biol 225:661–678. https://doi.org/10.1016/0022-2836(92)90393-X

176. Cranert S, Heyse S, Linger BR et al (2014) Tetrahymena Pot2 is a developmentally regulated paralog of Pot1 that localizes to chromosome breakage sites but not to telomeres. Eukaryot Cell 13:1519–1529. https://doi.org/10.1128/ec.00204-14

177. Fuhrmann G, Jönsson F, Weil PP et al (2016) RNA-template dependent de novo telomere addition. RNA Biol 13:733–739. https://doi.org/10.1080/15476286.2015.1134414

178. Greslin AF, Prescott DM, Oka Y et al (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. https://doi.org/10.1073/pnas.86.16.6264

179. Coyne RS, Stover NA, Miao W (2012) Whole Genome studies of Tetrahymena. Methods Cell Biol 109:53–81. https://doi.org/10.1016/B978-0-12-385967-9.00004-9

180. Prescott DM, Prescott JD, Prescott RM (2002) Coding properties of macronuclear DNA molecules in Sterkiella nova (Oxytricha nova). Protist 153:71–77. https://doi.org/10.1078/1434-4610-00084

181. Cervantes MD, Hamilton EP, Xiong J et al (2013) Selecting one of several mating types through gene segment joining and deletion in Tetrahymena thermophila. PLoS Biol 13:e1002284. https://doi.org/10.1371/journal.pbio.1001518

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.