Minireview

Transposons that clean up after themselves

Douglas L Chalker

Address: Biology Department, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130, USA.
Email: dchalker@biology2.wustl.edu

Published: 15 June 2009

*Genome **Biology*** 2009, **10:**224 (doi:10.1186/gb-2009-10-6-224)

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2009/10/6/224

Abstract

A transposon in the germline genome of the ciliate *Oxytricha* uses its transposase to remove itself, as well as other germline-limited DNA, from the differentiating somatic genome during development.

The genomes of eukaryotes are littered with transposon-derived sequences. As much as 45% of the human genome is composed of various types of mobile DNA elements or their remnants [1], and in plants such as maize, more than 60% of the genome consists of such repetitive sequences [2]. The ability of transposons to take over large chunks of genomes has given them the reputation of being selfish elements only interested in their own propagation. The fact that movement of these sequences into genes can inactivate genes and cause disease [3] is further evidence that they are up to no good. But can sequences that make up such a large fraction of genomes be all bad?

It is quite possible that mobile DNAs have positive influences on genomes that outweigh their deleterious effects [4]. Transposons do provide novel sequences that can be exploited by their host cells. Several examples are known of transposon functions that have been 'domesticated'; that is, a once-foreign sequence has been co-opted to carry out a host process (reviewed in [5]). For example, what was once probably some element's transposase, the enzyme that mediates the movement into and out of DNA, is now the mammalian Rag1/2 recombinase. This domestication event spurred the innovation of immunoglobulin gene rearrangement and the advent of the vertebrate adaptive immune system. The Rag1/2 proteins recognize recombination signal sequences (RSS), probably also derived from the transposon, that border hundreds of variable exons. Excision of the intervening sequences between selected exons leads to the assembly of a unique immunoglobulin gene from a non-expressible precursor locus [6]. This recombinase can be made to perform the reverse reaction and insert RSS-bound DNA into a target sequence, thereby acting as a transposase and supporting its likely transposon origin [7,8]. In a paper published recently in *Science*, Laura Landweber and colleagues (Nowacki *et al*. [9]) investigate another instance of the domestication of a transposon to carry out functional DNA rearrangements, and demonstrate the role of transposases in the extensive genome remodeling that produces the somatic genome in the ciliate *Oxytricha trifallax*.

Such benefits of transposable elements are realized over evolutionary time scales, so host organisms still need to manage their presence during an individual life span to keep any deleterious effects at minimum. Most cells do this by packaging regions of chromosomes with a high density of these elements into silent heterochromatin. By keeping these mobile elements in a transcriptionally silent state, they move relatively infrequently and have limited opportunities to cause deleterious mutations.

Ciliate genome remodeling

Ciliated protozoa such as *O. trifallax* appear to have worked out a unique solution to keep transposable elements in check: they eliminate them altogether from their transcriptionally active somatic nuclei. This solution is enabled by this organism's unique nuclear dimorphism (Figure 1a). *Oxytricha* contains both germline and somatic copies of its genome housed in separate nuclei called micro- and macronuclei, respectively, which have vastly different chromosome structures. The diploid micronucleus contains chromosomes of fairly typical size for a eukaryote, whereas the polyploid macronucleus contains aptly named 'nanochromosomes', whose average size is about 2 kbp (Figure 1a). Most of the more than 20,000 different macronuclear chromosomes

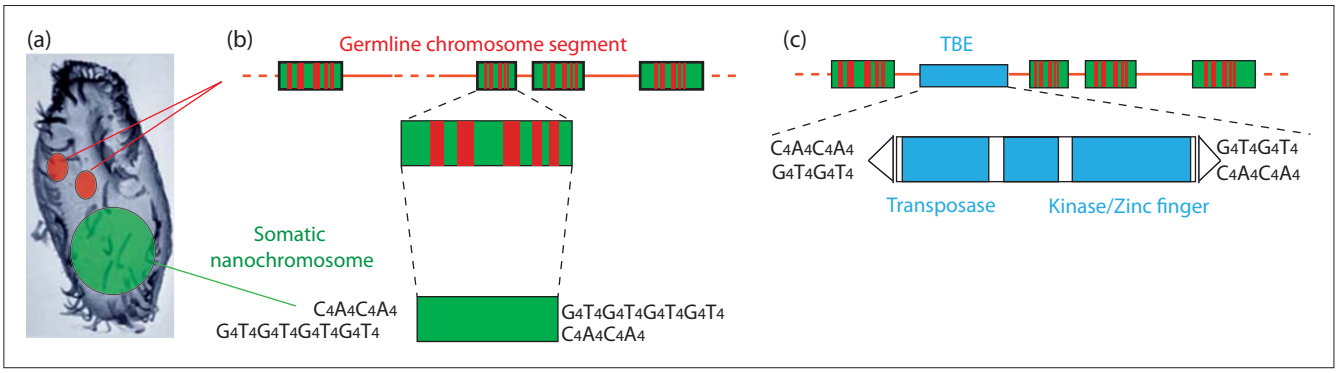

Figure 1
Oxytricha trifallax maintains functionally distinct genomes. **(a)** The ciliate *O. trifallax*. **(b)** The germline micronuclei (red circles) have chromosomes with large blocks of intergenic sequences (red lines) and coding sequences interrupted by IESs (red bars). During differentiation of the macronucleus (green circle), these intergenic sequences and IESs are removed from the large micronuclear chromosomes and the whole genome is fragmented into gene-size nanochromosomes. The ends of the nanochromosomes are stabilized by *de novo* addition of short telomeric sequences composed of G_4T_4 repeats. **(c)** A germline chromosome segment harboring a TBE element (blue box) in an intergenic region that will be eliminated during macronuclear differentiation. An enlargement of the TBE structure reveals the 20 bp of telomeric sequence flanking the element. Triangles represent 78-bp inverted repeats and the large blue boxes represent the three ORFs, including the transposase, that are under purifying selection [12].

contain a single gene flanked by short telomere sequences that together represent only 5% of the original germline sequence complexity [10].

The approximately 95% of the germline-limited DNA that is eliminated during the generation of the nanochromosomes is excised as either large intergenic sequence blocks or as short (tens to hundreds of base pairs in length) intragenic DNA segments called internal eliminated sequences (IESs). The transposons are largely eliminated along with the blocks of intergenic sequence, but some copies are eliminated from intragenic locations as well [11]. The short IESs number in the tens of thousands and have been postulated to be the remnants of transposable elements, but have diverged to the point that they are no longer recognizable as transposon-derived sequences. It has been something of a mystery how so many diverse sequences (transposons and IESs that share little or no similarity) can be coordinately excised from the developing somatic genome.

The recent work by Nowaki *et al.* [9] is an important step towards understanding this mystery. Their study demonstrates that a transposase encoded only in the germline genome plays a critical role in the genome-wide remodeling that produces the *Oxytricha* somatic genome. When the researchers used RNA interference (RNAi) to knock down the expression of the transposase encoded by the abundant germline telomere-bearing elements (TBEs), a type of DNA transposon, during macronuclear development, they found that a large fraction of the TBE elements failed to be eliminated from the genome as they should. Thus, the TBE transposase not only functions to move TBEs to new sites in the germline genome, but also eliminates these sequences from the somatic genome.

There are an estimated 2,000 copies of TBEs in the *Oxytricha* germline (Figure 1b). Their name derives from short stretches of telomeric repeat sequence (G_4T_4) that flank the inverted repeat ends of the integrated elements. Three divergent families of TBEs have been identified by comparing the predicted amino acid sequences of multiple transposase clones, and Landweber and colleagues had to knock down the transposase expression of all three to see a significant retention of TBEs in the somatic nucleus. This suggests that the transposases from all three types can mobilize any of the others. Some years ago it was found that the transposase, as well as the other two open reading frames (ORFs) of these elements, appears to be under purifying selection to preserve function [12]. That observation long spurred speculation that the transposase might play a role in the programmed genome rearrangements of the host. The new finding of Nowacki *et al.* [9] finally gives experimental support to that idea.

A particularly revealing result of this study is that the RNAi knock down of the transposase did not just interfere with the elimination of TBEs, but also reduced the efficiency of excision of the short IESs [9]. This was not easily predictable, as most IESs share little sequence similarity with TBEs. It is easy to envisage how a 'cut and paste' transposon such as a TBE could eliminate itself from the genome. These transposons move by the transposase recognizing terminal sequences of the transposon, and mediating both excision from the DNA and insertion in a new target site. The elimination process would be the same as transposition except that no new target locus is attacked - cut but not paste. The short IESs do not have TBE-like terminal sequences, and so it is unlikely that the transposase can recognize these DNA segments purely by identifying a conserved sequence.

Fortunately, a major clue to the mystery of how the many thousands of short (some as small as 14 bp), non-conserved IESs are initially identified was provided by a previous study from Landweber and colleagues in which they discovered noncoding transcripts that function as RNA guides for genome remodeling [13]. These RNA guides appear to be copies of the nanochromosomes (which lack all IESs) transcribed from the parental macronucleus during nuclear development. These noncoding RNAs are transported from the parental macronucleus into the undifferentiated somatic nucleus that contains the intact germline genome. In that work [13], the authors even introduced aberrantly rearranged RNA copies and produced rearranged chromosomes that modeled the mutant molecule. The implication of this result is that the noncoding RNAs not only identify the DNA that should be retained in the new macronucleus, but actually serve as a type of template that programs the new somatic genome.

Identification of the RNA guides was remarkable, but it still provided only a few clues as to how homologous RNAs might direct chromosomal rearrangements. Characterization of DNA molecules templated by introduced RNAs revealed that mismatches between the RNAs and the chromosome could be retained after rearrangement, a finding that argues that some RNA-directed DNA repair must have occurred. Such a mechanism would probably require that the DNA undergoing rearrangement must be nicked or cleaved to initiate repair. The TBE transposase is now the clear candidate to direct this repair-mediated DNA rearrangement. It is still not obvious how the RNA guides might recruit this transposase, so like any good mystery there are surely more twists in the plot to come before the ultimate solution is revealed.

The programmed DNA rearrangements of ciliates are these cells' means of protecting their genomes from the action of transposons; that is, transposons cannot spread if they are not expressed, although there must be windows during development when they are expressed and can expand in the germline. Whereas the RNA guides used in *Oxytricha* seem at first glance like a unique innovation, the mechanism used in distantly related ciliates such as *Paramecium* and *Tetrahymena* to carry out genome remodeling has striking parallels with the PIWI-interacting RNA (piRNA) pathway used to silence transposons in the metazoan germline [14]. Both these ciliates use small RNAs and a RNAi-related mechanism to guide their DNA rearrangements [15-17]. Like *Oxytricha*, these ciliates compare the parental rearranged genome with the undifferentiated genome in the developing macronucleus to help select which sequences should be eliminated [18,19]. The comparison of parental and developing genomes is probably mediated by the interaction of germline-produced small RNAs and noncoding transcripts from the macronucleus [20-22]. The comparison of genomes using RNA is not likely to be a ciliate-specific phenomenon, as recent studies have shown that the piRNAs can monitor compatibility between male and female genomes in *Drosophila* [23,24].

Domestication or mutualism?

It is quite intriguing that *Oxytricha* seems to have tamed the TBE transposon, using it to completely remodel its genome. This is not the first example of cells domesticating a transposon protein to keep mobile DNAs in check. The fission yeast *Schizosaccharomyces pombe* uses Cenp-b-like proteins, which are related to proteins from *pogo*-like elements, to silence its RNA transposons [25]. Do other ciliates, such as *Paramecium* and *Tetrahymena*, harbor abundant transposons in their germline that could provide the machinery to carry out developmental genome rearrangements? No obvious candidate elements have yet been found. A recent study of transcriptome profiling in *Tetrahymena* identified a transposase-like protein encoded in the macronuclear genome that is expressed exclusively during the time that IESs are excised [26]. If this proves to be the *Tetrahymena* IES excisase, this example looks more like a classical transposon domestication such as Rag1/2 or Cenp-b.

The taming of the TBE elements is hardly a typical example of transposon domestication. Instead, *Oxytricha* and its TBEs seem to have reached a mutual understanding. As Nowacki *et al*. [9] indicate, they can persist in the germline genome as long as they clean themselves out of the somatic nucleus.

Acknowledgements

DLC is supported by research grants from the National Institutes of Health (R01-GM069593) and the National Science Foundation (MCB 0642162).

References

1. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, *et al.*: **Initial sequencing and analysis of the human genome.** *Nature* 2001, **409:**860-921.
2. Messing J, Dooner HK: **Organization and variability of the maize genome.** *Curr Opin Plant Biol* 2006, **9:**157-163.
3. Steinmeyer K, Klocke R, Ortland C, Gronemeier M, Jockusch H, Grunder S, Jentsch TJ: **Inactivation of muscle chloride channel by transposon insertion in myotonic mice.** *Nature* 1991, **354:**304-308.
4. Kazazian HH Jr: **Mobile elements: drivers of genome evolution.** *Science* 2004, **303:**1626-1632.
5. Volff JN: **Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes.** *BioEssays* 2006, **28:**913-922.
6. Bassing CH, Swat W, Alt FW: **The mechanism and regulation of chromosomal V(D)J recombination.** *Cell* 2002, **109 (Suppl):**S45-S55.
7. Hiom K, Melek M, Gellart M: **DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations.** *Cell* 1998, **94:**463-470.
8. Agrawal A, Eastman QM, Schatz DG: **Transposition mediatd by RAG1 and RAG2 and its implications for the evolution of the immune system.** *Nature* 1998, **394:**744-751.

9. Nowacki M, Higgins BP, Maquilan GM, Swart EC, Doak TG, Landweber LF: **A functional role for transposases in a large eukaryotic genome.** *Science* 2009, **324:**935-938.
10. Prescott DM: **The DNA of ciliated protozoa.** *Microbiol Rev* 1994, **58:**233-267.
11. Williams K, Doak TG, Herrick G: **Developmental precise excision of *Oxytricha trifallax* telomere-bearing elements and formation of circles closed by a copy of the flanking target duplication.** *EMBO J* 1993, **12:**4593-4601.
12. Witherspoon DJ, Doak TG, Williams KR, Seegmiller A, Seger J, Herrick G: **Selection on the protein-coding genes of the TBE1 family of transposable elements in the ciliates *Oxytricha fallax* and *O. trifallax*.** *Mol Biol Evol* 1997, **14:**696-706.
13. Nowacki M, Vijayan V, Zhou Y, Schotanus K, Doak TG, Landweber LF: **RNA-mediated epigenetic programming of a genome-rearrangement pathway.** *Nature* 2008, **451:**153-158.
14. Malone CD, Hannon GJ: **Small RNAs as guardians of the genome.** *Cell* 2009, **136:**656-668.
15. Mochizuki K, Fine NA, Fujisawa T, Gorovsky MA: **Analysis of a *piwi*-related gene implicates small RNAs in genome rearrangement in tetrahymena.** *Cell* 2002, **110:**689-699.
16. Garnier O, Serrano V, Duharcourt S, Meyer E: **RNA-mediated programming of developmental genome rearrangements in *Paramecium tetraurelia*.** *Mol Cell Biol* 2004, **24:**7370-7379.
17. Lepere G, Nowacki M, Serrano V, Gout JF, Guglielmi G, Duharcourt S, Meyer E: **Silencing-associated and meiosis-specific small RNA pathways in *Paramecium tetraurelia*.** *Nucleic Acids Res* 2009, **37:**903-915.
18. Duharcourt S, Butler A, Meyer E: **Epigenetic self-regulation of developmental excision of an internal eliminated sequence in *Paramecium tetraurelia*.** *Genes Dev* 1995, **9:**2065-2077.
19. Chalker DL, Yao M-C: **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* 1996, **16:**3658-3667.
20. Chalker DL, Fuller P, Yao MC: **Communication between parental and developing genomes during tetrahymena nuclear differentiation is likely mediated by homologous RNAs.** *Genetics* 2005, **169:**149-160.
21. Lepere G, Betermier M, Meyer E, Duharcourt S: **Maternal noncoding transcripts antagonize the targeting of DNA elimination by scanRNAs in *Paramecium tetraurelia*.** *Genes Dev* 2008, **22:**1501-1512.
22. Aronica L, Bednenko J, Noto T, DeSouza LV, Siu KW, Loidl J, Pearlman RE, Gorovsky MA, Mochizuki K: **Study of an RNA helicase implicates small RNA-noncoding RNA interactions in programmed DNA elimination in *Tetrahymena*.** *Genes Dev* 2008, **22:**2228-2241.
23. Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ: **An epigenetic role for maternally inherited piRNAs in transposon silencing.** *Science* 2008, **322:**1387-1392.
24. Chambeyron S, Popkova A, Payen-Groschene G, Brun C, Laouini D, Pelisson A, Bucheton A: **piRNA-mediated nuclear accumulation of retrotransposon transcripts in the *Drosophila* female germline.** *Proc Natl Acad Sci USA* 2008, **105:**14964-14969.
25. Cam HP, Noma K, Ebina H, Levin HL, Grewal SI: **Host genome surveillance for retrotransposons by transposon-derived proteins.** *Nature* 2008, **451:**431-436.
26. Miao W, Xiong J, Bowen J, Wang W, Liu Y, Braguinets O, Grigull J, Pearlman RE, Orias E, Gorovsky MA: **Microarray analyses of gene expression during the *Tetrahymena thermophila* life cycle.** *PLoS ONE* 2009, **4:**e4429.