Reprint Series
24 June 1994, Volume 264, pp. 1870–1871

SCIENCE

RNA Editing and the Evolution of Parasites

Larry Simpson and Dmitri A. Maslov

RNA Editing and the Evolution of Parasites

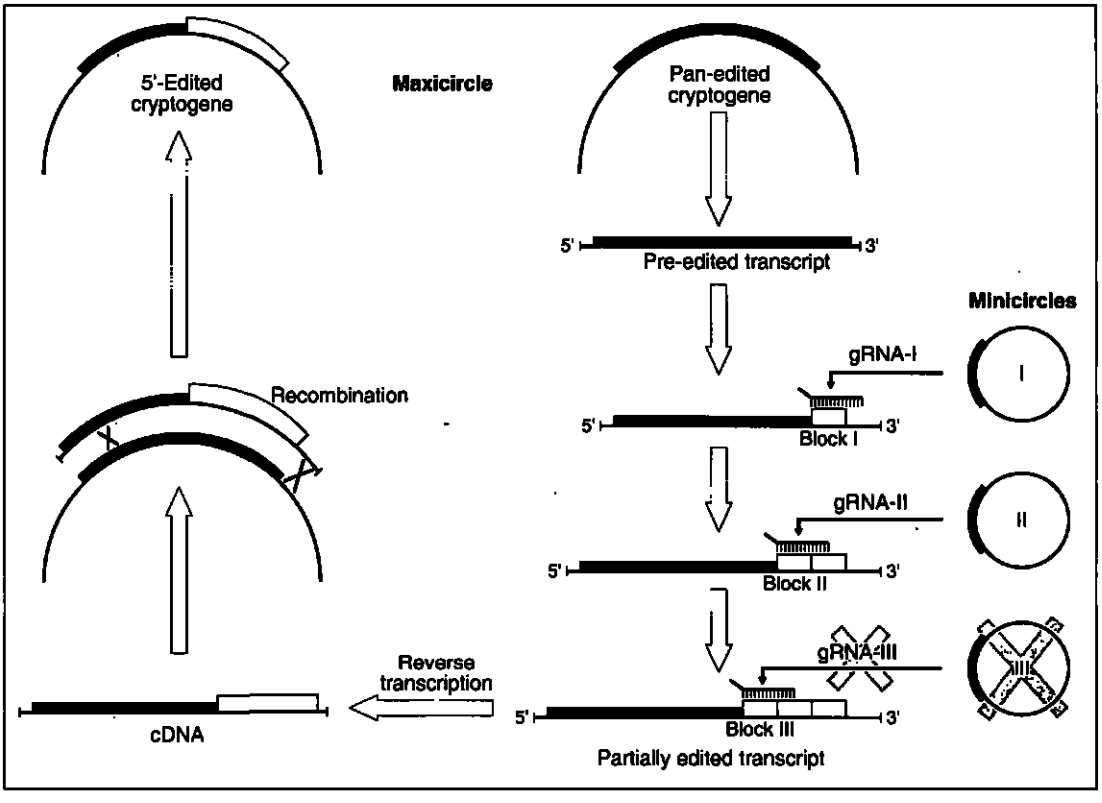
Larry Simpson and Dmitri A. Maslov

The kinetoplastid flagellates, together with their sister group of euglenoids, represent the earliest extant lineage of eukaryotic organisms containing mitochondria (*1*). Within the kinetoplastids, there are two major groups, the poorly studied bodonids-cryptobiids, which consist of both free-living and parasitic cells, and the better known trypanosomatids, which are obligate parasites (*2*).

Perhaps because of the antiquity of the trypanosomatid lineage, these cells possess several unique genetic features (see accompanying Perspective by Nilsen)—one of which is RNA editing of mitochondrial transcripts. This RNA editing function (*3–7*) creates open reading frames in "cryptogenes" by insertion (or occasional deletion) of uridine (U) residues at a few specific sites within the coding region of an mRNA (5'-editing) or at multiple specific sites throughout the mRNA (pan-editing). The precise number and location of U's to be inserted or deleted is determined by base-pairing with A or G residues in specialized RNAs called gRNAs. The mechanism of RNA editing is still an unsettled question, with transesterification or cleavage-ligation representing two alternative models (*8, 9*). The process occurs in ribonucleoprotein complexes containing gRNAs, mRNAs, and proteins.

The evolutionary origin of RNA editing in trypanosome mitochondria and its relationship to parasitism is of great interest. Was RNA editing an adaptation to parasitism, or an ancient genetic trait that was successfully used for this purpose? In both classical hypotheses for the origin of kinetoplastids, free-living bodo-like organisms were ancestral, but there is disagreement on the nature of the primary parasitic host. The "invertebrate first" model (*10, 11*) states that the initial parasitism was in the gut of pre-Cambrian invertebrates. Coevolution of parasite and host would have led to a wide distribution of trypanosomatids in insects and leeches. In this theory, digenetic life cycles (alternating invertebrate and vertebrate hosts) evolved later as a result of the acquisition by some hemipterans and dipterans of the ability to feed on the blood (hematophagy) of vertebrates. This hypothesis predicts that the monogenetic parasites of invertebrates would constitute the earliest diverging branches of the phylogenetic tree and that the digenetic parasites would have evolved later. The "vertebrate first" hypothesis (*12, 13*) states that parasitism first occurred in the gut of vertebrates and then in their blood. The evolution of leeches and hematophagous arthropods led to the appearance of a digenetic life cycle. Monogenetic parasites would represent cells that secondarily lost the ability to live in the vertebrate host.

In line with the "invertebrate first" hypothesis, constructs of an unrooted tree from mitochondrial ribosomal RNA (rRNA) sequences were arbitrarily rooted in the *Crithidia* branch (*14*). However, more recent phylogenetic reconstructions with nuclear rRNA sequences and with *Euglena* as an outgroup yielded a tree with an identical topology, but with a root in the *Trypanosoma* branch (*15–17*). In this tree, the bodonid-cryptobiid clade, represented by the free-living *Bodo caudatus* and the fish parasite *Trypanoplasma borreli*, constitutes an early diverging sister group to the trypanosomatids. In the trypanosomatid branch, digenetic organisms do not form a separate clade, suggesting either several independent losses of the digenetic life cycle, as suggested by Landweber and Gilbert (*16*), or several independent origins, as suggested by Fernandes and colleagues (*15*). The time of divergence of the trypanosomatid lineage can be estimated from the rRNA data to be approximately the time of appearance of vertebrates (*15*). In spite of the fact that the most deeply diverged ex-



Evolution of RNA editing in kinetoplastid protozoa. The primary transcript (thick black line) is edited by the first three overlapping gRNAs. Edited sequences, open boxes. The cDNA for the partially edited transcript replaces the original cryptogene in one of the maxicircles by homologous crossing-over. If the minicircle class encoding one of the three gRNAs is lost, cells lacking the substituted cryptogene could not edit this transcript, and this may be lethal. Cells with a substituted cryptogene would have a selective advantage.

The authors are in the Howard Hughes Medical Institute, Department of Biology, and Department of Medical Microbiology and Immunology, University of California at Los Angeles, Los Angeles, CA 90024, USA.

tant branch is represented by the digenetic trypanosomes, hematophagous invertebrate vectors appeared much later (*18*), suggesting that digeneity in the trypanosomes is a derived trait. However, since digeneity is a derived character in both classical models, it is clear that the molecular phylogenetic results have not resolved the origin of parasitism in the kinetoplastids.

A solution to this problem may lie in establishing the evolutionary relationships between the trypanosomes of fish, amphibians, and reptiles. If parasitism was first established in vertebrates, the parasites of vertebrates should form the most ancient lineages. On the other hand, if parasitism was first established in early invertebrates and the parasites were later inherited by insects from which digeneity arose, the leptomonad trypanosomatids that are found in invertebrates other than arthropods form the most ancient lineages.

Did the origin of RNA editing precede the origin of parasitism? A comparison of the extent of editing in homologous cryptogenes in different species yields the surprising result that pan-editing is a primitive evolutionary feature, and moderate or 5'-editing is a derived feature in the trypanosomatid lineage (*16*, *17*). Furthermore, our recent discovery of pan-editing in the parasitic cryptobiid *T. borreli* (*19*) pushes pan-editing and editing itself back in time to an ancestor of the entire kinetoplastid order. This would suggest that RNA editing may have preceded the appearance of an obligate parasitic life cycle. An ancient origin of pan-editing in the kinetoplastid lineage makes a primordial origin of a U- insertion or -deletion type of editing more plausible. However, it is still impossible to rule out a later origin (*7*, *20*) of U-insertion or -deletion editing within the early eukaryotic ancestors of kinetoplastids.

How did RNA editing evolve? Several times during the evolution of the kinetoplastids, pan-edited cryptogenes were substituted with less edited counterparts (*17*, *21*). Because editing proceeds 3' to 5' within an editing domain (*22–24*), the 5'-edited genes resemble the structures of partially edited mRNAs transcribed from a pan-edited cryptogene. The cryptogene substitutions could have involved complementary DNAs (cDNAs) derived from reverse transcription of partially edited mRNAs (*7*, *16*, *17*, *25*) (see figure), a mechanism that resembles that previously proposed for intron removal in yeast by RNA-mediated homologous recombination (*26*, *27*). The driving force for selection of such a retroposed copy could be the loss of one or more gRNA genes required for editing of the 3' region. Most gRNAs are encoded in catenated minicircle DNA molecules, which consist of multiple sequence classes varying in relative abundance. The loss of an entire class of low-copy number minicircle sequences (*24*) by missegregation at division of the kinetoplast (*28*) or by transkinetoplastidy (*29*) could cause the loss of a specific gRNA family. Cells with a retroposed, partially edited RNA would survive the loss of a specific gRNA gene family because the U's added by editing would already be genomically encoded. This phenomenon appears, in the case of the COIII gene of *Blastocrithidia culicis*, to have resulted in the complete replacement of an entire pan-edited gene (*17*), raising the possibility that all mitochondrial genes in the ancestral kinetoplastids were represented by (G + A)–rich skeletons, with the uridines being encoded by complementary A or G residues in multiple overlapping gRNA molecules.

Loss of minicircle classes occurs both in culture and in nature. A laboratory strain of *Leishmania tarentolae* contains 31 different gRNAs, 17 of which are encoded in minicircles (*24*). In constrast, a recently isolated wild strain of *L. tarentolae* contains 37 additional minicircle-encoded gRNAs, which mediate the editing of at least five additional cryptogenes (*30*), suggesting that a loss of minicircle sequence classes for apparently nonessential genes can occur during prolonged cultivation. A similar phenomenon occurs in nature when cyclical transmission is disrupted, as in the case of *Trypanosoma equiperdum*, a variant of *T. brucei* transmitted venereally, or in *Trypanosoma evansi*, a variant that is transmitted by the tabanid fly but does not propagate in this host. In these species the minicircle DNA has lost all sequence diversity and consists of a single sequence class (*31*, *32*), and the maxicircle DNA either has large deletions (*T. equiperdum*) or is absent (*T. evansi*) (*33*, *34*).

Editing appears to be such a labile genetic trait that it is indeed surprising that it still exists at all. This suggests that editing may have a selective advantage. In *T. brucei*, which undergoes reversible repression of mitochondrial synthetic activity during its digenetic life cycle, the editing of several genes is developmentally regulated (*6*), thereby possibly affecting mitochondrial translation by controlling the abundance of functional mRNAs. The extremely large minicircle and gRNA repertoire in this organism (over 900 different gRNAs) (*35*) may have evolved to provide a gRNA redundancy, such that the loss of a particular minicircle sequence class encoding a specific subset of gRNAs would not disrupt the editing cascade. In monogenetic kinetoplastids, the requirements for regulation of specific genes could be less stringent or absent, as a result of a simpler life cycle.

RNA editing seems to be an early evolutionary invention that came on the scene before the appearance of parasitism. Editing may have been inherited from the RNA world or developed within the early mitochondria in response to unknown regulatory demands. In the course of evolution, editing was partially or completely eliminated in many lineages. We hypothesize that it turned out to be useful for the development of parasitic adaptations, as exemplified by the developmental regulation of editing in *T. brucei*. Further understanding of the maintenance of RNA editing during the evolution of the kinetoplastid protozoa requires an understanding of its actual role in the life cycle, a problem for the future.

References and Notes

1. M. Sogin, J. Gunderson, H. Elwood, R. Alonso, D. Peattie, *Science* **243**, 75 (1989).
2. K. Vickerman, in *Biology of the Kinetoplastida*, W. H. R. Lumsden and D. A. Evans, Eds. (Academic Press, New York, 1976), pp. 1–34.
3. R. Benne *et al.*, *Cell* **46**, 819 (1986).
4. R. Benne, *Mol. Biol. Rep.* **16**, 217 (1992).
5. S. L. Hajduk, M. E. Harris, V. W. Pollard, *FASEB J.* **7**, 54 (1993).
6. K. Stuart, in *RNA Editing—The Alteration of Protein Coding Sequences of RNA*, R. Benne, Ed. (Horwood, New York, 1993), pp. 25–52.
7. L. Simpson, D. A. Maslov, B. Blum, in (6), pp. 53–85.
8. T. R. Cech, *Cell* **64**, 667 (1991).
9. B. Blum, N. R. Sturm, A. M. Simpson, L. Simpson, *ibid.* **65**, 543 (1991).
10. L. Leger, *Comptes Rendus de Seances de la Societe de Biologie et de Ses Filiales* **56**, 615 (1904).
11. J. R. Baker, *Parasitic Protozoa* (Hutchinson, London, 1973).
12. E. A. Minchin, *Q. J. Microsc. Sci.* **52**, 159 (1908).
13. F. G. Wallace, *Exp. Parasitol.* **18**, 124 (1966).
14. J. Lake *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4779 (1988).
15. A. P. Fernandes *et al.*, *ibid.* **90**, 11608 (1993).
16. L. F. Landweber and W. Gilbert, *ibid.* **91**, 918 (1994).
17. D. A. Maslov, H. A. Avila, J. A. Lake, L. Simpson, *Nature* **365**, 345 (1994).
18. F. Lambrecht, *Proc. Am. Philos. Soc.* **124**, 367 (1980).
19. D. A. Maslov and L. Simpson, unpublished data.
20. P. S. Covello and M. W. Gray, *Trends Genet.* **9**, 265 (1993).
21. L. F. Landweber and W. Gilbert, *Nature* **363**, 179 (1993).
22. J. Abraham, J. Feagin, K. Stuart, *Cell* **55**, 267 (1988).
23. N. R. Sturm and L. Simpson, *ibid.* **61**, 871 (1990).
24. D. A. Maslov and L. Simpson, *ibid.* **70**, 459 (1992).
25. L. F. Landweber, *BioSystems* **28**, 41 (1992).
26. G. R. Fink, *Cell* **49**, 5 (1987).
27. L. D. Derr and J. Strathern, *Nature* **361**, 170 (1993).
28. P. Borst, *Trends. Genet.* **7**, 139 (1991).
29. S. T. Lee *et al.*, *Mol. Cell. Biol.* **14**, 587 (1994).
30. O. Thiemann, D. A. Maslov, L. Simpson, unpublished data.
31. Z.-R. Lun, R. Brun, W. Gibson, *Mol. Biochem. Parasitol.* **50**, 189 (1992).
32. H.-H. Shu and K. Stuart, *Nucleic Acids Res.* **21**, 2951 (1993).
33. A. Frasch *et al.*, *Biochim. Biophys. Acta* **607**, 397 (1980).
34. P. Borst, F. Fase-Fowler, W. Gibson, *Mol. Biochem. Parasitol.* **23**, 31 (1987).
35. R. A. Corell *et al.*, *Nucleic Acids Res.* **21**, 4313 (1993).

36. We thank O. Thiemann for assistance with the illustration.