Curr Genet (2011) 57:225–232
DOI 10.1007/s00294-011-0340-8

REVIEW

Gene fragmentation: a key to mitochondrial genome evolution in Euglenozoa?

Pavel Flegontov · Michael W. Gray · Gertraud Burger · Julius Lukeš

Received: 6 April 2011 / Accepted: 8 April 2011 / Published online: 5 May 2011

Abstract Phylum Euglenozoa comprises three groups of eukaryotic microbes (kinetoplastids, diplonemids, and euglenids), the mitochondrial (mt) genomes of which exhibit radically different modes of organization and expression. Gene fragmentation is a striking feature of both euglenid and diplonemid mtDNAs. To rationalize the emergence of these highly divergent mtDNA types and the existence of insertion/deletion RNA editing (in kinetoplastids) and trans-splicing (in diplonemids), we propose that in the mitochondrion of the common evolutionary ancestor of Euglenozoa, small expressed gene fragments promoted a rampant neutral evolutionary pathway. Interactions between small antisense transcripts of these gene fragments and full-length transcripts, assisted by RNA-processing enzymes, permitted the emergence of RNA editing and/or trans-splicing activities, allowing the system to tolerate indel mutations and further gene fragmentation, respectively, and leading to accumulation of additional mutations. In this way, dramatically different mitochondrial genome structures and RNA-processing machineries were able to evolve. The paradigm of constructive neutral evolution acting on the widely different mitochondrial genetic systems in Euglenozoa posits the accretion of initially neutral molecular interactions by genetic drift, leading inevitably to the observed 'irremediable complexity'.

Keywords Euglena · Diplonema · Mitochondrial genome · RNA editing · Constructive neutral evolution

Communicated by H. Jacobs.

P. Flegontov · J. Lukeš (✉)
Biology Centre, Institute of Parasitology,
Czech Academy of Sciences, and Faculty of Sciences,
University of South Bohemia, 37005 České Budějovice,
Czech Republic
e-mail: jula@paru.cas.cz

M. W. Gray
Centre for Comparative Genomics and Evolutionary Bioinformatics, and Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, NS B3H 4R2, Canada

G. Burger
Robert-Cedergren Center in Bioinformatics and Genomics, and Department of Biochemistry, Université de Montréal, Montreal, QC H3C 3J7, Canada

Introduction

Kinetoplastid flagellates contain an astonishingly complex mitochondrial (mt) DNA network, known since the 1950s and termed kinetoplast (k) DNA. This intricate assemblage of interconnected circular DNA molecules of different sizes, along with the more recently discovered and extremely sophisticated uridine (U) insertion/deletion editing of kDNA transcripts, stretches the imagination of molecular and evolutionary biologists. Very recent studies of mtDNA and mtRNA structure and processing from diplonemids and euglenids, which are specifically related to kinetoplastids, have revealed dramatically different complexities, for which a logical evolutionary pathway from a common ancestor of these three protist groups is hard to imagine. Yet in the light of these discoveries, a recent incarnation (Gray et al. 2010; Lukeš et al. 2011) of a neutral evolutionary theory (Covello and Gray 1993) termed 'constructive neutral evolution' (Stoltzfus 1999) becomes particularly appealing. In this paper, we argue that if kDNA structure and the kinetoplastid type of RNA editing are perhaps singular in detail, what is not unique is their almost nonsensical molecular complexity.

The proposed eukaryotic supergroup Excavata (Hampl et al. 2009) brings together unicellular eukaryotic microbes (protists) in which the evolutionary fate of mtDNA has been markedly different. Whereas the parabasalids (metamonads) *Trichomonas* and *Giardia* have apparently lost all DNA from their highly modified mitochondria, the jakobid flagellate *Reclinomonas*, the malawimonad *Malawimonas* and the heterolobosean *Naegleria* have retained the most gene-rich mitochondrial genomes known (Gray et al. 2004). Finally, Euglenozoa, which represents another prominent phylum within this supergroup, features mitochondrial genomes with reduced gene content, but with complex architectures and modes of expression. Remarkably, various groups of the free-living or parasitic members of Excavata evolved dramatically different mitochondrial genomes, culminating in a diversity that is virtually unmatched in any other eukaryotic lineage.

The phylum Euglenozoa is subdivided into three well-supported groups: Kinetoplastida, Diplonemida, and Euglenida (Moreira et al. 2004; Simpson and Roger 2004). The nuclear and mitochondrial genomes of kinetoplastids have been studied in great detail, primarily because this taxon includes the medically important trypanosomatids (Lukeš et al. 2005); in marked contrast, very few investigations have focused on the mitochondrial genomes of the other two euglenozoan groups, with the mtDNA of diplonemids (Kiethega et al. 2011; Marande et al. 2005; Marande and Burger 2007; Roy et al. 2007; Vlček et al. 2011) and euglenids (Roy et al. 2007; Spencer and Gray 2011) only recently receiving some attention.

Mitochondrial DNA in Euglenozoa

The most prominent feature of the mitochondrial genetic system of kinetoplastids is the extensive editing of mRNAs. Small guide (g) RNAs and several large protein complexes mediate the insertion and/or deletion of U residues at numerous pre-ordained sites within most mitochondrial transcripts. The intricacies of this mechanism are reflected in the organelle's genetic structure, with protein- and rRNA-coding genes residing in so-called maxicircles, which are the counterpart of typical mitochondrial genomes, while the vast majority of the gRNA genes are located in so-called minicircles (for review see Lukeš et al. 2005). Usually, several gRNAs are needed to re-tailor a single transcript, so that thousands of mutually interlocked minicircles are found to code for hundreds of different gRNAs, with a species-specific number of gRNA genes per minicircle (for review see Liu et al. 2005). While essentially the same type of editing seems to be present in all kinetoplastids (Deschamps et al. 2011), kDNA is variably organized and packaged. The ancestral state, found in various bodonids, is represented by a morphologically dispersed kinetoplast composed of non-catenated circles, which may constitute up to half of the total cellular DNA (Lukeš et al. 1998). Catenation of individual DNA circles into a single kDNA network, a feature common to all obligatorily parasitic trypanosomatids (Liu et al. 2005), allows more faithful replication leading to a highly compact and much smaller kinetoplast (Lukeš et al. 2002).

Diplonemids, the sister group of kinetoplastids (Simpson and Roger 2004), have evolved in a different direction. The gene content of their mitochondrial genome is very similar to that of kinetoplastids, but the mtDNA is not organized as maxi- and minicircles (Marande et al. 2005). All genes appear to be fragmented, with each gene fragment encoded by a single DNA circle (Marande and Burger 2007; Vlček et al. 2011). These subgenic modules are independently transcribed and then trans-spliced by an as yet unknown mechanism. Neither conserved elements, intron signatures nor sequence complementarity could be identified at the ends of individual transcript modules (Kiethega et al. 2011). How the transcript fragments are correctly spliced is a challenging question, and it has been hypothesized that as yet unidentified gRNA-like molecules might be mediators of trans-splicing. Remarkably, some form of RNA editing exists in parallel with trans-splicing: insertion of six U residues between two modules of the *cox1* transcript is observed (Kiethega et al. 2011; Marande and Burger 2007). Trans-splicing and RNA editing, as well as a multi-partite mitochondrial genome consisting of small, circular, repeat-rich chromosomes of two size classes, have been documented in all four species investigated to date: *Diplonema ambulator*, *D. papillatum*, *Diplonema* sp.*2*, and *Rhynchopus euleeides* (Kiethega et al. 2011).

Mitochondrial genomes of euglenids have remained intractable for a long time. The mitochondrial genome of the hallmark species *Euglena gracilis* is represented by a heterogeneous distribution of linear, ~4 kb-long molecules (Spencer and Gray 2011). To date, the genes *cox1*, *cox2,* and *cox3*, not requiring transcript editing, as well as SSU and LSU rRNAs that are each split into two fragments, have been identified in *Euglena* mitochondria (Spencer and Gray 2011; Tessier et al. 1997; Yasuhira and Simpson 1997). Genes are flanked by complex arrays of direct and inverted repeats; overall, genome organization appears chaotic, with full-length genes, gene fragments, repeats, and repeat clusters appearing in various arrangements. The most conspicuous feature of this genome is that small dispersed gene fragments retain almost perfect identity with sequences in the corresponding full-length authentic genes.

Peranema trichophorum is a basally branching non-photosynthetic euglenid (Adl et al. 2005). Electron

microscopy indicates that its mtDNA consists of linear molecules of a broad size range between 1 and 75 kb, while restriction enzyme digestion reveals a band pattern similar to that observed in *E. gracilis* (Roy et al. 2007). Yet another non-photosynthetic and basally branching euglenid, *Petalomonas cantuscygni*, exhibits a quite different mitochondrial genome structure. Electron microscopy, restriction enzyme digestion, and DNA sequencing of the bona fide mtDNA indicate a 40 kbp circle (Roy et al. 2007). The failure to identify in this genome typical mitochondrial genes suggests the possibility of massive gene fragmentation and/or transcript editing of an unknown type.

Constructive neutral evolution model

An obvious question follows: how could this array of 'strange' and at the same time strikingly different mitochondrial genomes in Euglenozoa evolve from conventional gene-rich mtDNAs preserved in the sister lineages Jakobida and Heterolobosea?

In recent years, alternatives to the exclusively adaptationist evolutionary approach favored by many molecular biologists have appeared (Koonin 2009). There is a growing understanding that mutation bias, drift and selectively neutral 'ratchets', rather than positive selection, may underpin the main evolutionary mechanisms for genomes, multi-molecular machines, and regulatory networks (Covello and Gray 1993; Gray et al. 2010; Lukeš et al. 2011; Lynch 2007; Lynch and Conery 2003; Stoltzfus 1999). We propose that the model of constructive neutral evolution (Covello and Gray 1993; Stoltzfus 1999), recently expanded as an explanation for the seemingly 'irremediable complexity' of the molecular systems and machines characteristic of eukaryotes (Gray et al. 2010; Lukeš et al. 2011), is particularly well suited to explain the explosion of forms and complexities of mitochondrial genomes and their expression in euglenozoans. This model assumes that fortuitous interactions between proteins and/or nucleic acids relax functional constraints on some mutations that would otherwise be deleterious or even lethal, so that mutations arising in one component (say A) can become fixed in the genome by genetic drift as a result of the prior interaction of A with another component (say B). In other words, the initially neutral interaction of A with its partner B 'pre-adapts' A for subsequent mutations, by effectively 'pre-suppressing' these mutations, thereby rendering A functionally dependent on B. Through additional mutations, the system's complexity increases, and because there are more ways to build a complex system than to build a simple one, and because a random walk through 'complexity space' can continue indefinitely, molecular structures evolve that are unnecessarily complex and seemingly 'sophisticated' in function (Gray et al. 2010; Lukeš et al. 2011). It is evident that such a ratchet can function in an essentially neutral manner, initially without positive selection. Nevertheless, complex systems that emerge in this way can be recruited subsequently for secondary functions under positive selection (Covello and Gray 1993).

While overall genome architecture has not often been a subject of adaptationist interpretations, RNA editing has received much more attention. This phenomenon has been viewed as a relict of the RNA world (Benne 1990; Stuart et al. 1997), as an error-correction mechanism (Stuart et al. 1997; Cavalier-Smith 1997), or on the contrary, as a mutagenic mechanism (Landweber and Gilbert 1993). Lifestyle with changing requirements for mitochondrial function has been viewed as a prerequisite for the origin of U insertion/deletion RNA editing in kinetoplastids. The idea here is that under anaerobic conditions, mutations that accumulate in non-expressed genes in maxicircle DNA might be effectively suppressed via RNA-processing activities that are recruited when these genes are expressed under aerobic conditions (Cavalier-Smith 1997). Speijer (2006, 2007, 2010) hypothesized that gRNAs, dispersed over multiple genomic locations and needed for editing a few essential transcripts, hinder large-scale deletions in kDNA at life cycle stages when other transcripts are non-essential. Furthermore, roles for RNA editing in regulation of gene expression (Stuart et al. 1997) and generation of protein diversity (Ochsenreiter and Hajduk 2006, 2007; Ochsenreiter et al. 2008) have been proposed. However, even from an adaptationist perspective, all these benefits of RNA editing either appear marginal when compared to their costs, or they become useful only after the emergence of editing. In our view, none of these arguments convincingly accounts for the emergence of the editing system in the first place.

In the context of a constructive neutral evolutionary perspective, we propose that the emergence of editing in the mitochondrion of kinetoplastid flagellates proceeded through three stages (Covello and Gray 1993; Gray 2003; Horton and Landweber 2002; Keeling et al. 2010; Lukeš et al. 2009; Simpson and Maslov 1999; Stoltzfus 1999): (1) gene fragments appeared and, by chance, some of them produced antisense transcripts; (2) various enzymes—endonuclease, exonuclease, uridylyl transferase, RNA ligase, and RNA helicase, normally present in the mitochondrion as components of other RNA-processing pathways—fortuitously interacted with the antisense transcripts (=gRNA precursors); (3) these interactions allowed mutations in genes to be corrected at the transcript level, through the concerted action of the various editing enzymes together with gRNAs. Thus, mutations that would be deleterious or lethal in the absence of editing are maintained and spread by

genetic drift. Reversion to the original, non-edited state is unlikely given that additional mutations may arise within the region that is complementary to the initial antisense transcript (gRNA), and additional gRNAs may be recruited. Nevertheless, reversion of almost all mutations is conceivable if the edited transcript is reverse-transcribed and the resulting cDNA replaces the gene sequence whose transcript requires editing (Landweber and Gilbert 1993; Maslov et al. 1994; Simpson and Maslov 1999; Simpson et al. 2000).

Although accepting the three-step model, Simpson and Maslov (1999) proposed that random sequences, rather than gene fragments, could be recruited as gRNAs. Indeed, the short length of gRNAs and imperfect matches to edited mRNAs (featuring frequent G–U base pairing) make them nearly indistinguishable from random sequences in computer simulations (von Haeseler et al. 1992). There is one example in *Phytomonas* of a random sequence being recruited as a gRNA after loss of the original gRNA due to large deletions in the maxicircle (Nawathean and Maslov 2000). Expanding on the only known contemporary example of *cis* editing (in the transcript of the trypanosomatid *cox2* gene), Golden and Hajduk (2005) proposed that unedited transcripts themselves, i.e., their random fragments, were initially used in *cis* as gRNAs. There is no easy way to distinguish between the two alternatives (random vs. non-random) for the source of primordial gRNAs. Guide RNAs could originate from perfectly matching gene fragments, and then become more fragmented and accumulate mutation by genetic drift as long as G–U base pairs were tolerated by the editing machinery, resulting in sequences nearly indistinguishable from random ones. Selection for wobble pairing was proposed for gRNAs because weak G–U base pairs in the gRNA–mRNA duplex are readily displaced by the perfectly matching anchor region of the following gRNA, allowing progression of editing in the 3′–5′ direction along the template (Horton and Landweber 2002; Simpson and Maslov 1999).

Evolution of euglenozoan mitochondrial genomes

With the constructive neutral evolutionary scenario in mind and considering information recently obtained about mtDNA in diplonemids (Kiethega et al. 2011; Vlček et al. 2011) and *E. gracilis* (Spencer and Gray 2011), we are in a position to speculate more generally about mitochondrial genome evolution in Euglenozoa. The finding of small gene fragments in the mtDNA of *E. gracilis* has important evolutionary implications, suggesting that such fragments were present in the mtDNA of the euglenozoan last common ancestor (ELCA) and subsequently recruited as gRNA precursors along the diplonemid–kinetoplastid branch (Spencer and Gray 2011).

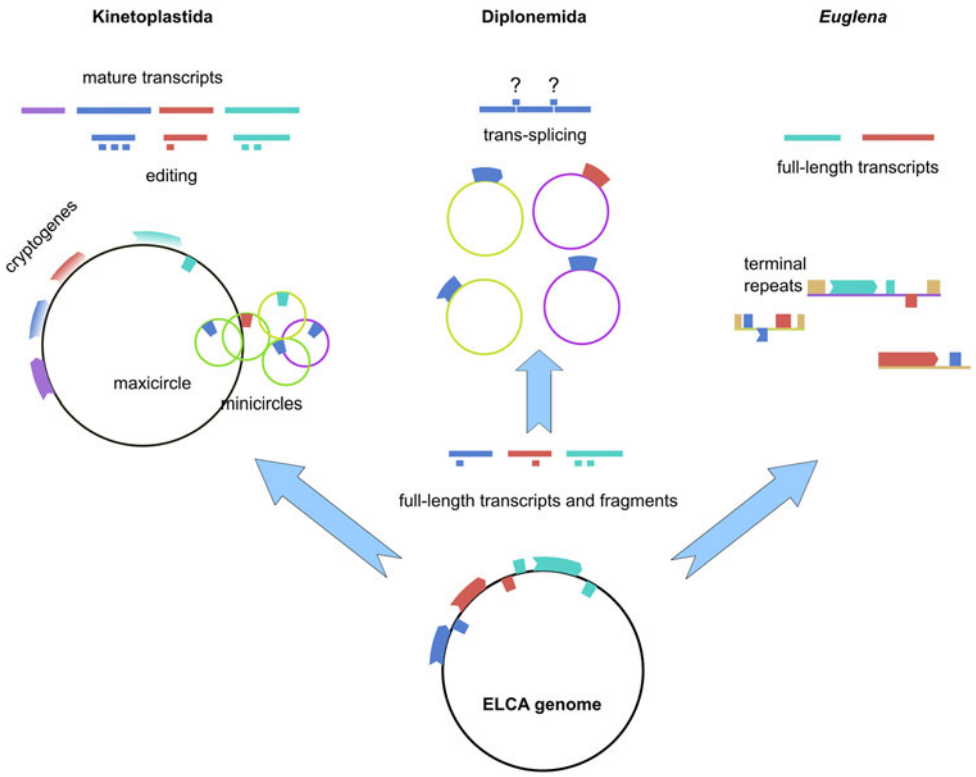

Fig. 1 Schematic representation of euglenozoan mitochondrial genomes (in trypanosomatids, diplonemids, and *Euglena*) and proposed genome organization in their euglenozoan last common ancestor (ELCA). The 5′ and 3′ ends of full-length genes are depicted by *arrowheads*, cryptogenes by gradient fill, and transcription direction by the position of the shape inside or outside of the *genomic circle*. Transcripts and processing events are shown above the *genomic circles*

So, what did the ELCA mitochondrial genome look like? What pathways led to the current striking diversity and complexity of mitochondrial genomes in Euglenozoa? We infer that the ELCA genome was probably circular because kinetoplastids [and apparently the euglenid *Petalomonas* (Roy et al. 2007)] still retain an unbroken main mitochondrial genome. But the ELCA genome already contained 'seeds of future Chaos'—gene fragments arising by recombination between short repeats, degradation of auto-spliced introns in gene duplicates, or simple pseudogenization (Fig. 1). It should be noted that to the extent that most euglenozoans have only a single mitochondrion per cell but multiple genome copies per organelle, all genes may be considered multi-copy, and so under somewhat relaxed selective pressure. Gene copies on separate molecules could take different evolutionary paths, e.g., some could become fragmented due to genome rearrangements.

With increasing recombination frequency, such a metastable genome might progress to a truly chaotic state such as that observed in *Euglena*. Chaos is most probably a one-way road, as it is extremely unlikely that a complete 'normal' genome could be reassembled from a highly fragmented and rearranged state. Notably, 'genomic chaos' also seems to be the rule in some other mtDNAs, such as those of dinoflagellates (Nash et al. 2008; Waller and Jackson 2009) and *Chromera velia*, the closest photosynthetic relative of Apicomplexa (Flegontov et al., submitted).

The constructive neutral evolution model implies that the transcribed small gene fragments could be recruited for different activities: for example, transcript trans-splicing in Diplonemida, leading to more and more widespread gene fragmentation and trans-splicing, or U insertion/deletion RNA editing in Kinetoplastida (although, as discussed earlier, random sequences could also, in principle, be recruited as gRNAs) (Fig. 1). We can speculate that the decision as to which pathway would be taken—editing, trans-splicing, both or yet some other process—would depend on which enzymatic machineries were most readily available. Both trans-splicing and RNA editing machineries seem to coexist in *D. papillatum*, although the involvement of gRNA-like molecules in these processes remains to be demonstrated in this case (Kiethega et al. 2011). It is notable that RNA editing has not been detected so far in *Euglena*, possibly because the observed gene fragments are not transcribed (Spencer and Gray 2011). Capping of triphosphorylated 5′ ends, a very sensitive method used to detect kinetoplastid gRNAs, has failed to reveal any small gRNA-like molecules in *E. gracilis* mitochondria (Yasuhira and Simpson 1997). However, an unknown type of editing probably exists in the basal-branching euglenid *P. cantuscygni* (Roy et al. 2007).

Notably, a role in RNA editing has been suggested for gene fragments in the extremely scrambled mtDNAs of dinoflagellates, as some fragments were found to encode edited positions (Nash et al. 2007). Editing in mitochondria of dinoflagellates is not as extensive as in kinetoplastids, yet it is widespread and of several different types (Lin et al. 2002; Nash et al. 2008; Waller and Jackson 2009). However, such edited fragments were found for only a minor fraction of edited sites, and their role remains unknown (Waller and Jackson 2009).

In kinetoplastids, most gRNA genes became segregated either on free or mutually interlocked minicircles (Liu et al. 2005; Lukeš et al. 1998), or arranged in large tandem arrays in megacircles, as in the case of *Trypanoplasma borreli* (Yasuhira and Simpson 1996). Regardless of their genomic location, minicircles could have been formed initially by recombination between direct repeats flanking gene fragments (Spencer and Gray 2011). However, evolution of the mitochondrial genome architecture found solely in diplonemids is more difficult to envisage. All circular chromosomes of *D. papillatum* share a 1.5-kb sequence stretch, and all chromosomes within a given class have a ~5 kb region in common, which includes conspicuous tandem repeat arrays (Marande and Burger 2007; Vlček et al. 2011).

An intriguing question is what determines an equilibrium level of editing or trans-splicing, and a more 'philosophical' question follows: are the current systems in equilibrium? A high rate of genetic drift (i.e., low effective population size) is required for the constructive neutral evolution ratchet to function efficiently. A relatively low mutation rate is also required as all complexities of gene structure and expression increase significantly the mutational target size for a gene (Lynch 2007). On the other hand, the 'complexity burden' can become too high and so counteract the ratchet. Such a theoretical complexity plateau might have been reached by diplonemids and kinetoplastids, which have an extremely high mtDNA content per cell (Marande et al. 2005; Roy et al. 2007; Lukeš et al. 1998), and whose editing and processing machineries require dozens or even hundreds of dedicated proteins, as has been demonstrated in the model flagellate *Trypanosoma brucei* (Lukeš et al. 2005). In *D. papillatum*, each gene is split into 3–12 fragments (Vlček et al. 2011). Further fragmentation might be deleterious due to the associated required increase in mtDNA content (each single coding module is embedded in a ~7 kb chromosome) or by the inability to splice certain junctions due to inherent limitations of the splicing machinery. The finding of much the same gene fragmentation across different diplonemid taxa (Kiethega et al. 2011) is suggestive of a complexity plateau.

In trypanosomatids, a tendency was noted for editing events to be more abundant in the basal genus *Trypanosoma* (Maslov et al. 1994) than in derived genera; in this

case, editing sites were likely lost by means of reverse transcription (Kolesnikov et al. 2003; Landweber and Gilbert 1993; Simpson and Maslov 1999; Simpson et al. 2000). *Trypanosoma* species have the most complex life cycles among trypanosomatids, and the correlation between life cycle complexity and the extent of RNA editing has received considerable attention in the form of sophisticated hypotheses (Speijer 2006, 2007, 2010, 2011). This latter author doubts that a neutral ratchet would be able to generate the extensive editing observed in kinetoplastids, and argues that although the initial occurrence of limited editing might follow the model of Covello and Gray (1993), rampant spread of editing, probably early in the kinetoplastid lineage, required some selective advantage (Speijer 2007). Speijer suggests that the most plausible advantage is a scrambled genome itself, because universally essential gRNA genes scattered in the genome prevent accumulation of large-scale deletions and chromosome losses at life cycle stages (or in certain environments) when most of the mitochondrial genes are dispensable. In this scenario, rampant accumulation of gRNA genes was selected for in lineages that evolved such metabolic versatility, the dixenous trypanosomes in particular.

However, there may be a rather simple explanation: complex life cycles in parasites imply severe population bottlenecks, leading to genetic drift and further spread of editing, which were counteracted by reverse transcription and substitution in species with simpler life cycles. It remains to be seen whether this view also holds true for the kinetoplastid groups emerging prior to the Trypanosomatidae, collectively called bodonids. Unfortunately, little information is currently available about editing patterns in these groups (Blom et al. 1998; Deschamps et al. 2011; Lukeš et al. 1994; Simpson and Maslov 1999; Yasuhira and Simpson 1996). Moreover, segregation of most gRNAs on separate genomic molecules (minicircles) can render the genome even more vulnerable to degradation at certain life cycle stages (Savill and Higgs 1999), especially because of the sheer bulk of DNA and selective disadvantage probably associated with it: partial or complete minicircle losses are extensively documented in *T. brucei* (Lai et al. 2008; Schnaufer et al. 2002, 2005) and *Leishmania tarentolae* (Gao et al. 2001; Thiemann et al. 1994). Thus, in our view, gRNA genes do not prevent this bulk of DNA from being lost, but it is selection at the species level that prevents *petite* (=dyskinetoplastic) strains, constantly generated de novo, from spreading, since they usually lack efficient transmission mechanisms (Lun et al. 2010; Schnaufer 2010).

The existence of at least one transcript whose editing is expanding throughout the evolution of kinetoplastids, considered an example of editing 'on-the-way-in' (Blom et al. 1998), does not necessarily contradict the view that in these protists the evolutionary trend is toward loss and not gain of editing sites. We interpret this example rather as a reflection of an equilibrium between the loss and emergence of editing. The fact that no kinetoplastid flagellate has yet been described in which either all of its mitochondrial mRNAs are edited, or none is, testifies to such an equilibrium. It is notable that some genes are never edited in any trypanosomatid species (Lukeš et al. 2005). In these cases, it is plausible that the respective transcribed fragments, or gRNA precursors, never emerged.

Conclusions

It was generally expected that the description of elusive mitochondrial genomes of diplonemids and especially euglenids would help us to understand the origin and evolution of one of the most complex genomic systems in nature, namely the mitochondrial kDNA network and the U insertion/deletion type of RNA editing in trypanosomes and other kinetoplastids. Contrary to expectations, diplonemid and euglenid genomes do not appear to represent any intermediary steps or key evolutionary events leading to the above-mentioned kinetoplastid machineries; nevertheless, their investigation has provided provocative evolutionary insights. We postulate that in the mitochondrion of the common evolutionary ancestor of *Euglena*, *Diplonema,* and *Trypanosoma*, gene fragmentation allowed the operation of rampant constructive neutral evolution, leading in each lineage to dramatically different DNA replication and RNA-processing pathways. The only 'unifying feature' of euglenozoan mitochondrial genomes is their individuality and baroque intricacy, which have been revealed so far by intense decade long studies only in the medically relevant trypanosomes. It is safe to predict that the biochemical mechanisms operating on mtDNA and mtRNA in diplonemids and euglenids may even surpass in complexity the biochemical mechanisms underlying kinetoplastid RNA editing and kDNA replication.

References

Adl S, Simpson AGB, Farmer MA, Andersen RA et al (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J Euk Microbiol 52:399–451

Benne R (1990) RNA editing in trypanosomes—is there a message? Trends Genet 6:177–181

Blom D, de Haan A, van den Berg M, Sloof P, Jirků M, Lukeš J, Benne R (1998) RNA editing in the free-living bodonid *Bodo saltans*. Nucl Acids Res 26:1205–1213

Cavalier-Smith T (1997) Cell and genome coevolution: facultative anaerobiosis, glycosomes and kinetoplastan RNA editing. TIG 13:6–9

Covello PS, Gray MW (1993) On the evolution of RNA editing. Trends Genet 9:265–268

Deschamps P, Lara E, Marande W, López-García P, Ekelund F, Moreira D (2011) Phylogenomic analysis of kinetoplastids supports that trypanosomatids arose from within bodonids. Mol Biol Evol 28:53–58

Flegontov P, Lai DH, Jirků M, Janouškovec J, Oborník M, Keeling PJ, Lukeš J (submitted) The mitochondrial genome of *Chromera velia* encodes a single conserved gene

Gao G, Kapushoc ST, Simpson AM, Thiemann OH, Simpson L (2001) Guide RNAs of the recently isolated LEM125 strain of *Leishmania tarentolae*: an unexpected complexity. RNA 7:1335–1347

Golden DE, Hajduk SL (2005) The 3′-untranslated region of cytochrome oxidase II mRNA functions in RNA editing of African trypanosomes exclusively as a cis guide RNA. RNA 11:29–37

Gray MW (2003) Diversity and evolution of mitochondrial RNA editing systems. IUBMB Life 55:227–233

Gray MW, Lang BF, Burger G (2004) Mitochondria of protists. Annu Rev Genet 38:477–524

Gray MW, Lukeš J, Archibald JM, Keeling PJ, Doolittle WF (2010) Irremediable complexity? Science 330:920–921

Hampl V, Hug L, Leigh JW, Dacks JB, Lang BF, Simpson AG, Roger AJ (2009) Phylogenetic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups". Proc Natl Acad Sci USA 106:3859–3864

Horton TL, Landweber LF (2002) Rewriting the information in DNA: RNA editing in kinetoplastids and myxomycetes. Curr Opin Microbiol 5:620–626

Keeling PJ, Leander BS, Lukes J (2010) Reply to Speijer et al.: Does complexity necessarily arise from selective advantage? Proc Natl Acad Sci USA 107:E26

Kiethega GN, Turcotte M, Burger G (2011) Evolutionary conserved cox1 trans-splicing without cis-motifs. Mol Biol Evol. doi:10.1093/molbev/msr075

Kolesnikov AA, Merzliak EM, Bessolitsyna EA, Fediakov AV, Schonian G (2003) Reduction of the edited domain of the mitochondrial A6 gene for ATPase subunit 6 in Trypanosomatidae. Mol Biol (Mosk) 37:637–642

Koonin EV (2009) Darwinian evolution in the light of genomics. Nucl Acids Res 37:1011–1034

Lai DH, Hashimi H, Lun ZR, Ayala FJ, Lukeš J (2008) Adaptations of *Trypanosoma brucei* to gradual loss of kinetoplast DNA: *Trypanosoma equiperdum* and *Trypanosoma evansi* are petite mutants of *T. brucei*. Proc Natl Acad Sci USA 105:1999–2004

Landweber LF, Gilbert W (1993) RNA editing as a source of genetic variation. Nature 363:179–182

Lin S, Zhang H, Spencer DF, Norman JE, Gray MW (2002) Widespread and extensive editing of mitochondrial mRNAs in dinoflagellates. J Mol Biol 320:727–739

Liu B, Liu Y, Motyka SA, Agbo EE, Englund PT (2005) Fellowship of the rings: the replication of kinetoplast DNA. Trends Parasitol 21:363–369

Lukeš J, Arts GJ, van den Burg J, de Haan A, Opperdoes F, Sloof P, Benne R (1994) Novel pattern of editing regions in mitochondrial transcripts of the cryptobiid *Trypanoplasma borreli*. EMBO J 13:5086–5098

Lukeš J, Jirků M, Avliyakulov N, Benada O (1998) Pankinetoplast DNA structure in a primitive bodonid flagellate, *Cryptobia helicis*. EMBO J 17:838–846

Lukeš J, Guilbride DL, Votýpka J, Zíková A, Benne R, Englund PT (2002) Kinetoplast DNA network: evolution of an improbable structure. Eukaryot Cell 1:495–502

Lukeš J, Hashimi H, Zíková A (2005) Unexplained complexity of the mitochondrial genome and transcriptome in kinetoplastid flagellates. Curr Genet 48:277–299

Lukeš J, Leander BS, Keeling PJ (2009) Cascades of convergent evolution: the corresponding evolutionary histories of euglenozoans and dinoflagellates. Proc Natl Acad Sci USA 106:9963–9970

Lukeš J, Archibald JM, Keeling PJ, Doolittle WF, Gray MW (2011) How a neutral evolutionary ratchet can build cellular complexity. IUBMB Life, accepted

Lun ZR, Lai DH, Li FJ, Lukeš J, Ayala FJ (2010) *Trypanosoma brucei*: two steps to spread out from Africa. Trends Parasitol 26:424–427

Lynch M (2007) The frailty of adaptive hypotheses for the origins of organismal complexity. Proc Natl Acad Sci USA 104:8597–8604

Lynch M, Conery JS (2003) The origins of genome complexity. Science 302:1401–1404

Marande W, Burger G (2007) Mitochondrial DNA as a genomic jigsaw puzzle. Science 318:415

Marande W, Lukeš J, Burger G (2005) Unique mitochondrial genome structure in diplonemids, the sister group of kinetoplastids. Eukaryot Cell 4:1137–1146

Maslov DA, Avila HA, Lake JA, Simpson L (1994) Evolution of RNA editing in kinetoplastid protozoa. Nature 368:345–348

Moreira D, López-García P, Vickerman K (2004) An updated view of kinetoplastid phylogeny using environmental sequences and a closer outgroup: proposal for a new classification of the class Kinetoplastea. Int J Syst Evol Microbiol 54:1861–1875

Nash EA, Barbrook AC, Edwards-Stuart RK, Bernhardt K, Howe CJ, Nisbet RE (2007) Organization of the mitochondrial genome in the dinoflagellate *Amphidinium carterae*. Mol Biol Evol 24:1528–1536

Nash EA, Nisbet RER, Barbrook AC, Howe CJ (2008) Dinoflagellates: a mitochondrial genome all at sea. Trends Genet 24:328–335

Nawathean P, Maslov DA (2000) The absence of genes for cytochrome c oxidase and reductase subunits in maxicircle kinetoplast DNA of the respiration-deficient plant trypanosomatid *Phytomonas serpens*. Curr Genet 38:95–103

Ochsenreiter T, Hajduk SL (2006) Alternative editing of cytochrome c oxidase III mRNA in trypanosome mitochondria generates protein diversity. EMBO Rep 7:1128–1133

Ochsenreiter T, Hajduk SL (2007) The function of RNA editing in trypanosomes. In: Goringer HU (ed) RNA editing. Springer-Verlag, Berlin-Heidelberg, pp 181–197

Ochsenreiter T, Cipriano M, Hajduk SL (2008) Alternative mRNA editing in trypanosomes is extensive and may contribute to mitochondrial protein diversity. PLoS One 3:e1566

Roy J, Faktorová D, Lukeš J, Burger G (2007) Unusual mitochondrial genome structures throughout the Euglenozoa. Protist 158:385–396

Savill NJ, Higgs PG (1999) A theoretical study of random segregation of minicircles in trypanosomatids. Proc Biol Sci 266:611–620

Schnaufer A (2010) Evolution of dyskinetoplastic trypanosomes: how, and how often? Trends Parasitol 26:557–558

Schnaufer A, Domingo GJ, Stuart K (2002) Natural and induced dyskinetoplastic trypanosomatids: how to live without mitochondrial DNA. Int J Parasitol 32:1071–1084

Schnaufer A, Clark-Walker GD, Steinberg AG, Stuart K (2005) The F1–ATP synthase complex in bloodstream stage trypanosomes has an unusual and essential function. EMBO J 24:4029–4040

Simpson L, Maslov DA (1999) Evolution of the U-insertion/deletion RNA editing in mitochondria of kinetoplastid protozoa. Ann NY Acad Sci 870:190–205

Simpson AG, Roger AJ (2004) Protein phylogenies robustly resolve the deep-level relationships within Euglenozoa. Mol Phylogenet Evol 30:201–212

Simpson L, Thiemann OH, Savill NJ, Alfonzo JD, Maslov DA (2000) Evolution of RNA editing in trypanosome mitochondria. Proc Natl Acad Sci USA 97:6986–6993

Speijer D (2006) Is kinetoplastid pan-editing the result of an evolutionary balancing act? IUBMB Life 58:91–96

Speijer D (2007) Evolutionary aspects of RNA editing. In: Goringer HU (ed) RNA editing. Springer-Verlag, Berlin-Heidelberg, pp 199–227

Speijer D (2010) Constructive neutral evolution cannot explain current kinetoplastid panediting patterns. Proc Natl Acad Sci USA 107:E25

Speijer D (2011) Does constructive neutral evolution play an important role in the origin of cellular complexity? Making sense of the origins and uses of biological complexity. BioEssays. doi:10.1002/bies.201100010

Spencer DF, Gray MW (2011) Ribosomal RNA genes in *Euglena gracilis* mitochondrial DNA: fragmented genes in a seemingly fragmented genome. Mol Genet Genomics 285:19–31

Stoltzfus A (1999) On the possibility of constructive neutral evolution. J Mol Evol 49:169–181

Stuart K, Allen TE, Heidmann S, Seiwert SD (1997) RNA editing in kinetoplastid protozoa. Microbiol Mol Biol Rev 61:105–120

Tessier LH, van der Speck H, Gualberto JM, Grienenberger JM (1997) The cox1 gene from *Euglena gracilis*: a protist mitochondrial gene without introns and genetic code modifications. Curr Genet 31:208–213

Thiemann OH, Maslov DA, Simpson L (1994) Disruption of RNA editing in *Leishmania tarentolae* by the loss of minicircle-encoded guide RNA genes. EMBO J 13:5689–5700

Vlček C, Marande W, Teijeiro S, Lukeš J, Burger G (2011) Systematically fragmented genes in a multipartite mitochondrial genome. Nucl Acids Res 39:979–988

von Haeseler A, Blum B, Simpson L, Sturm N, Waterman MS (1992) Computer methods for locating kinetoplastid cryptogenes. Nucl Acids Res 20:2717–2724

Waller RF, Jackson CJ (2009) Dinoflagellate mitochondrial genomes: stretching the rules of molecular biology. BioEssays 31:237–245

Yasuhira S, Simpson L (1996) Guide RNAs and guide RNA genes in the cryptobiid kinetoplastid protozoan, *Trypanoplasma borreli*. RNA 2:1153–1160

Yasuhira S, Simpson L (1997) Phylogenetic affinity of mitochondria of *Euglena gracilis* and kinetoplastids using cytochrome oxidase I and hsp60. J Mol Evol 44:341–347