Review

Dinoflagellates: a mitochondrial genome all at sea

Edmund A. Nash, R. Ellen R. Nisbet, Adrian C. Barbrook and Christopher J. Howe

Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, UK

Dinoflagellate algae are notorious for their highly unusual organization of nuclear and chloroplast genomes. Early studies on the dinoflagellate mitochondrial genome indicated that it encodes the same three protein-coding genes found in *Plasmodium* spp., but with a complex organization and transcript editing. Recent work has extended this view, showing that the dinoflagellate mitochondrial genome contains a wide array of gene fragments and genes interspersed with noncoding inverted repeats. The genome seems to require noncanonical start and stop codons, as well as high levels of editing, *trans*-splicing and the addition of oligonucleotide caps at the 5′ and 3′ ends of transcripts. Despite its small coding content, the dinoflagellate mitochondrial genome is one of the most complex known.

Mitochondria and their genomes

Mitochondria are the organelles classically responsible for oxidative phosphorylation in eukaryotes, although they carry out a wide range of other functions, including the biogenesis of iron-sulphur clusters and some other prosthetic groups [1]. They are the result of an ancient symbiosis between an aerobic prokaryote and an anaerobic host, and their origin might mark the birth of eukaryotes [2]. The anaerobic mitochondria of some metazoans, such as worms, typically contain a partial electron transport chain and ATP synthase. However, they use molecules such as fumarate or nitrate instead of oxygen as terminal electron acceptors [3]. In some anaerobic protozoa, the mitochondria have been modified into hydrogenosomes, which generate hydrogen as a by-product of pyruvate oxidation for ATP production [1,4]. In other anaerobic protozoa, the mitochondria have been trimmed back further to mitosomes, whose sole function might be in iron-sulphur cluster biogenesis [1,4].

Mitochondria retain a genome that is a remnant of their prokaryotic ancestor. A prokaryotic genome of a few thousand genes has been reduced a hundredfold via gene loss and transfer to the host nucleus. The mitochondrial genome has been lost in mitosomes and in most hydrogenosomes (with the exception of the anaerobic ciliate *Nyctotherus ovalis* [5] and the stramenopile *Blastocystis* spp. [6]). Mitochondrial genomes have a very wide size range, from a few kilobase pairs to >2 Mbp in cucurbit plants such as muskmelon and cucumber [7]. Although the mitochondrial genome is essential for biogenesis of the organelle, the coding content varies widely, from three protein-coding genes in the malaria parasite *Plasmodium* spp. (see below) to just <100 protein coding genes in the jakobid flagellate *Reclinomonas americana* [8]. However, the coding content of the mitochondrial genome does not correlate directly with size. The *Reclinomonas* genome is ~69 kb, whereas the *Arabidopsis thaliana* mitochondrial genome contains only 32 protein-coding genes in 367 kb [9]. Molecules encoded in mitochondrial genomes commonly include subunits of the NADH dehydrogenase complex, succinate dehydrogenase, the cytochrome bc_1 complex, cytochrome oxidase and ATP synthase, as well as components of a gene expression machinery [including large and small ribosomal RNAs (rRNAs), a 5S rRNA, and transfer RNAs (tRNAs)] and less commonly, mitochondrial protein translocation machinery (reviewed in Ref. [10]). These are only a fraction of the complete set of mitochondrial proteins. The genes for the remainder now reside in the nucleus, with their products imported into the mitochondrion after synthesis.

There is a particular diversity of mitochondrial genomes in the large protist group known as the alveolates. This group comprises the Apicomplexa (which include many parasitic organisms such as *Plasmodium* spp., *Theileria* spp. and *Toxoplasma* spp.), the ciliates (such as *Tetrahymena* spp. and *Paramecium* spp.) and the dinoflagellates (see Figure II in Box 1). Dinoflagellates are unicellular algae that inhabit marine or brackish environments, and about one half of all known species are photosynthetic [11]. Photosynthetic dinoflagellates are ecologically important as major primary producers, as critical components of coral reefs [12] and as the cause of toxic 'red tides' [13].

Studies of small regions of the mitochondrial genomes of dinoflagellate algae over the last few years suggested a very unusual genome organization [14–19]. Recently, more systematic surveys in several species have indicated that dinoflagellate mitochondrial genomes are even more unusual than expected. The novelties include a highly complex genome structure, the presence of many pseudogenes and partial gene fragments, exceptionally diverse transcript editing and a dense pattern of inverted repeats within the DNA [20–24].

Mitochondrial DNA in the alveolate protists

Apicomplexa have the smallest coding content known of all mitochondrial genomes, containing just three protein-coding genes [25]. These comprise *cob*, encoding the cytochrome *b* polypeptide of the cytochrome bc_1 complex (complex III) of the electron transfer chain, and *cox1* and *cox3*, encoding two subunits of cytochrome oxidase (complex IV) [25,26]. In *Plasmodium* spp., these three genes are

Corresponding author: Howe, C.J. (c.j.howe@bioc.cam.ac.uk).

0168-9525/$ – see front matter © 2008 Elsevier Ltd. All rights reserved. doi:10.1016/j.tig.2008.04.001 Available online 29 May 2008

Box 1. Unusual features of dinoflagellate nuclear and chloroplast genomes

The first indication of an unusual nuclear genome in dinoflagellates came from ultrastructural studies that observed permanently condensed chromosomes throughout the cell cycle (Figure I). Dinoflagellate nuclear genomes can be very large, with estimates from 3 to 200 pg/cell in some species (reviewed in Ref. [55]). Dinoflagellates lack conventional histones, but a protein has been identified with some similarity to HU, a basic DNA binding protein found in eubacteria [56]. The unique organization of the dinoflagellate nucleus was at one time termed 'mesokaryotic', reflecting the view that the dinoflagellates represented a particularly primitive eukaryotic group with a nuclear complexity midway between that of prokaryotes and eukaryotes (discussed in Ref. [57]). However, phylogenetic analyses have indicated that dinoflagellates represent a highly derived rather than a basal eukaryotic lineage and that histone loss occurred secondarily [58]. Such phylogenies place the dinoflagellates within the alveolate protists, and as sisters of the Apicomplexa (Figure II).

Dinoflagellate chloroplast genomes are no less unusual. Despite early inferences of a conventionally sized genome in dinoflagellates [59], it became clear that chloroplast genes adopt a unique arrangement, with a series of plasmid-like minicircles each containing from one to a few genes [60,61]. Chloroplast genomes of plants and algae typically contain ~100 genes, so the complement of 14 protein-coding genes in dinoflagellates is highly reduced and makes up the smallest known chloroplast genome [36]. Chloroplast genes are transcribed polycistronically to form a single transcript from each minicircle. These transcripts are cleaved to gene-sized pieces, with intergenic regions being degraded [42,62].

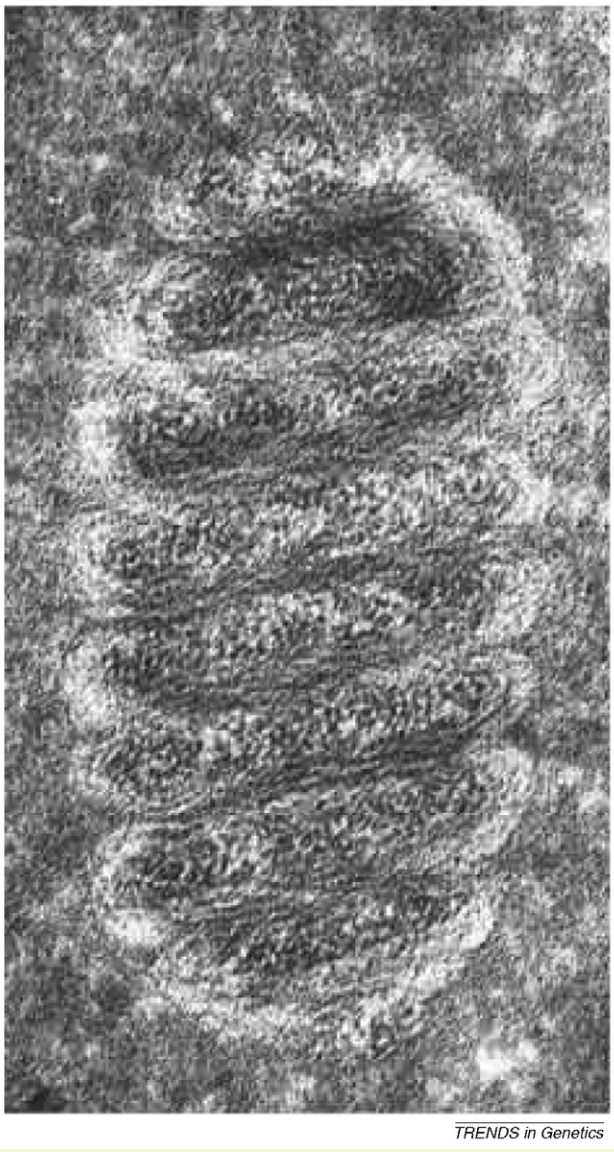

Figure I. A dinoflagellate chromosome. Single condensed chromosome in the nucleus of *Crypthecodinium cohnii* (reproduced with permission of The Company of Biologists from Ref. [63]; page 1234, Figure 2C).

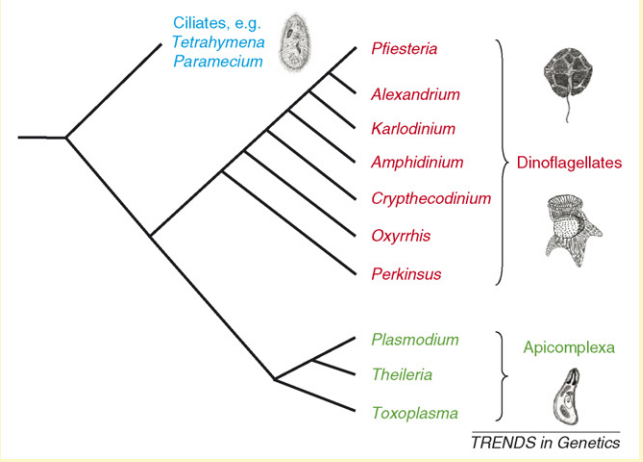

Figure II. Alveolate protists. Simplified diagram depicting some of the more well-known alveolates and the proposed evolutionary relationship of the lineages [30,64]. Line drawings were kindly provided by Dr. Giselle Walker.

organized as part of a repeating 6-kb unit. Apicomplexan mitochondrial DNA (mtDNA) also contains many sequences encoding small fragments of mitochondrial large and small rRNAs, which are located on both strands of the mtDNA (reviewed in Ref. [26]). It is unclear exactly how apicomplexan mtDNA replicates, although a complex recombination-mediated process using circular and branched intermediates is likely to be involved [27].

Ciliate mtDNA consists of a circular mapping molecule containing ~50 genes [28]. Unusual features of ciliate mtDNA include a split *nad1* gene (encoding subunit 1 of NADH dehydrogenase of complex I) with the two parts located on opposite DNA strands. The ciliate mitochondrial genome also contains split large subunit rRNA genes and highly divergent protein coding regions, including an extra sequence in *cox1* that produces a protein containing a unique ~100 amino acid insertion [28,29].

Because the dinoflagellate and apicomplexan lineages are thought to be sister groups (to the exclusion of the ciliates [30]; see Figure II in Box 1), it might be expected that dinoflagellate mtDNA would resemble the compact linear units of *Plasmodium* spp. However, the first glimpses of the dinoflagellate mitochondrial genome made a decade ago [14] yielded very confusing information. In retrospect, this is perhaps not very surprising, because the dinoflagellates also display an unconventional organization of their nuclear and chloroplast genomes (Box 1). In the last year, however, several publications have revealed a more detailed

picture of the structure and complexity of dinoflagellate mtDNA across several species [20–24].

Surprising complexity of dinoflagellate mtDNA

The first gene identified in dinoflagellate mtDNA was *cox1* in *Crypthecodinium cohnii* [14]. However, clones containing the gene showed different 5′ flanking sequences and an internal probe hybridized to several restriction fragments of mtDNA, some of which appeared to be present in a higher copy number than others. Subsequent analysis indicated the existence of distinct *cox1* coding sequences, possibly generated by recombination, placing different flanking sequences around a common central region [15]. Studies of *Pfiesteria* spp. demonstrated the existence of *cox3* and *cob* genes in mtDNA, but with a range of different intergenic regions and a variably sized *cob* coding region [17,18]. By contrast, Northern analysis of RNA from *C. cohnii* showed *cox1* transcripts of a single size [15]. Chaput *et al*. [16] showed *cox3* cDNAs from the dinoflagellate *Gonyaulax polyedra* had a range of sizes. They also found chimeric transcripts containing regions of *cox1* and *cob* as well as *cox3*. Another study indicated the existence of a large number of inverted repeats close to the *C. cohnii cox1* sequence [15]. Interestingly, analyses of *cob* sequences from six species indicated the occurrence of editing, primarily A→G (or A→I), U→C and C→U [19].

Complex organization of three protein coding genes

Recent work has confirmed and extended these indications of an unusual mitochondrial genome organization in dinoflagellates. The species investigated include *Alexandrium catenella* [20], *Amphidinium carterae* [21], *C. cohnii* [23], and *Karlodinium micrum* [23] as well as *Oxyrrhis marina* [22], which diverged early from the other dinoflagellate lineages [31], and *Durinskia baltica* and *Kryptoperidinium foliaceum*, which contain mitochondrial genomes from both a dinoflagellate host and a diatom endosymbiont [24]. Mitochondrial sequence data were obtained using primarily two approaches. One approach isolated an AT-rich satellite band above nuclear DNA in caesium chloride density gradients, and sequenced clones produced from this fraction picked either at random or by hybridization to probes for genes expected to reside within mtDNA. The other approach used PCR with primers based on expressed sequence tag (EST) sequences for probablemitochondrial genes. In all cases, only the same three protein-coding sequences, *cox1*, *cox3* and *cob*, were found (apart from rRNA, as discussed below). The sequences obtained varied from fragments of coding regions to potentially complete coding regions (setting aside uncertainties over start and stop codons, as discussed below). As had been observed in earlier studies, the flanking sequences of many of the potential genes and gene fragments found in any one strain

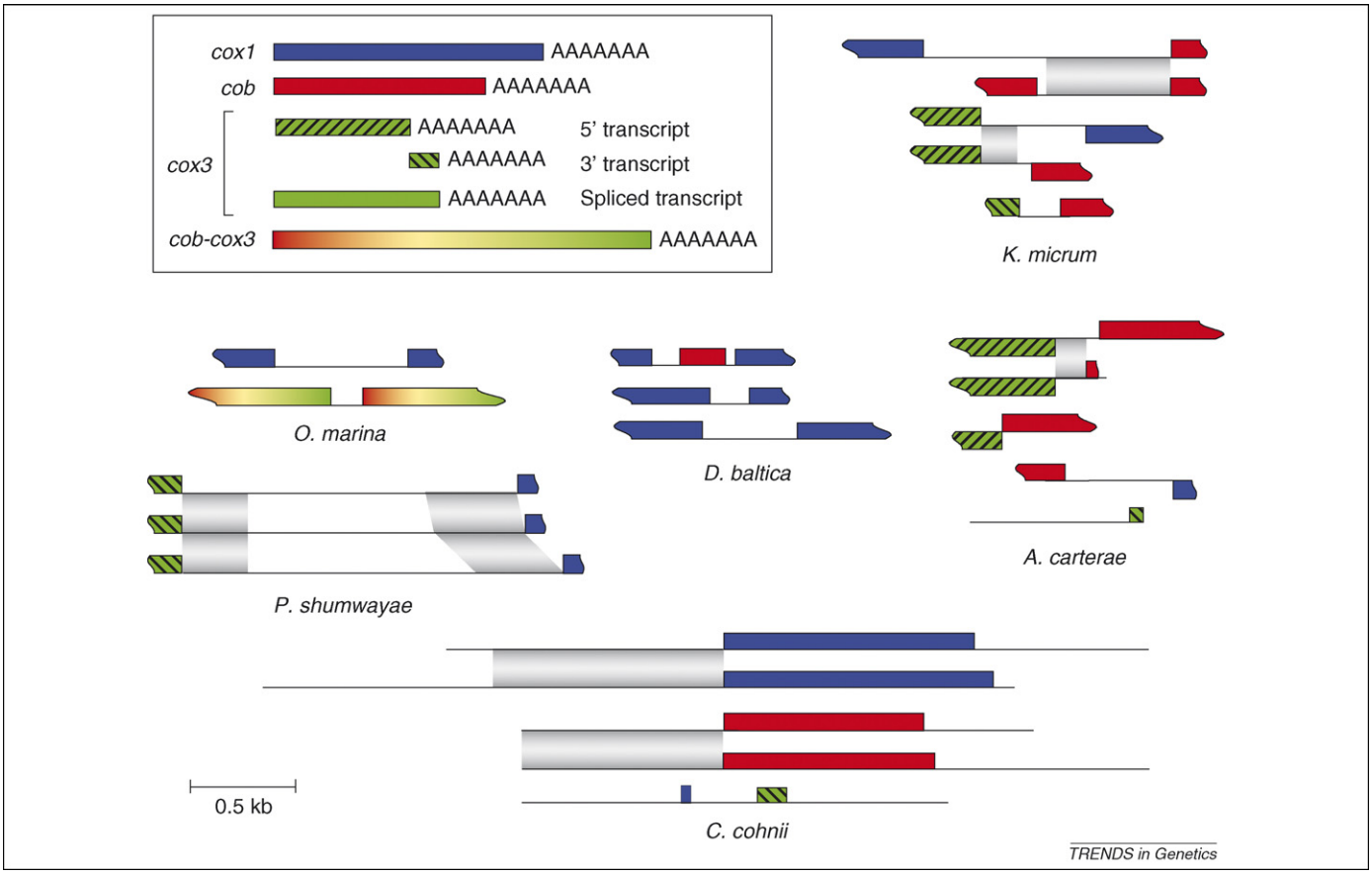

Figure 1. Arrangements of dinoflagellate mitochondrial DNA (mtDNA). Schematic diagram depicting various arrangements of *cox1*, *cox3* and *cob* sequences in six dinoflagellate species. Transcripts of the three genes (boxed) are shown to allow length comparisons. Ribosomal DNA (rDNA) sequences are not shown. Gray shading indicates homologous regions of noncoding DNA. A vertical line at an end of a coding region indicates the end of recognizable conserved coding sequence and not necessarily the 5′ or 3′ end of the open reading frame. A curved line at one end of a coding region indicates that the ends of the clone are delineated by PCR primers. The gene arrangements shown are examples and do not comprise a comprehensive list; see Refs. [17,21–24] for full details. *cox1*, blue; *cob*, red; *cox3*, green; *cob-cox3*, transition from red to green.

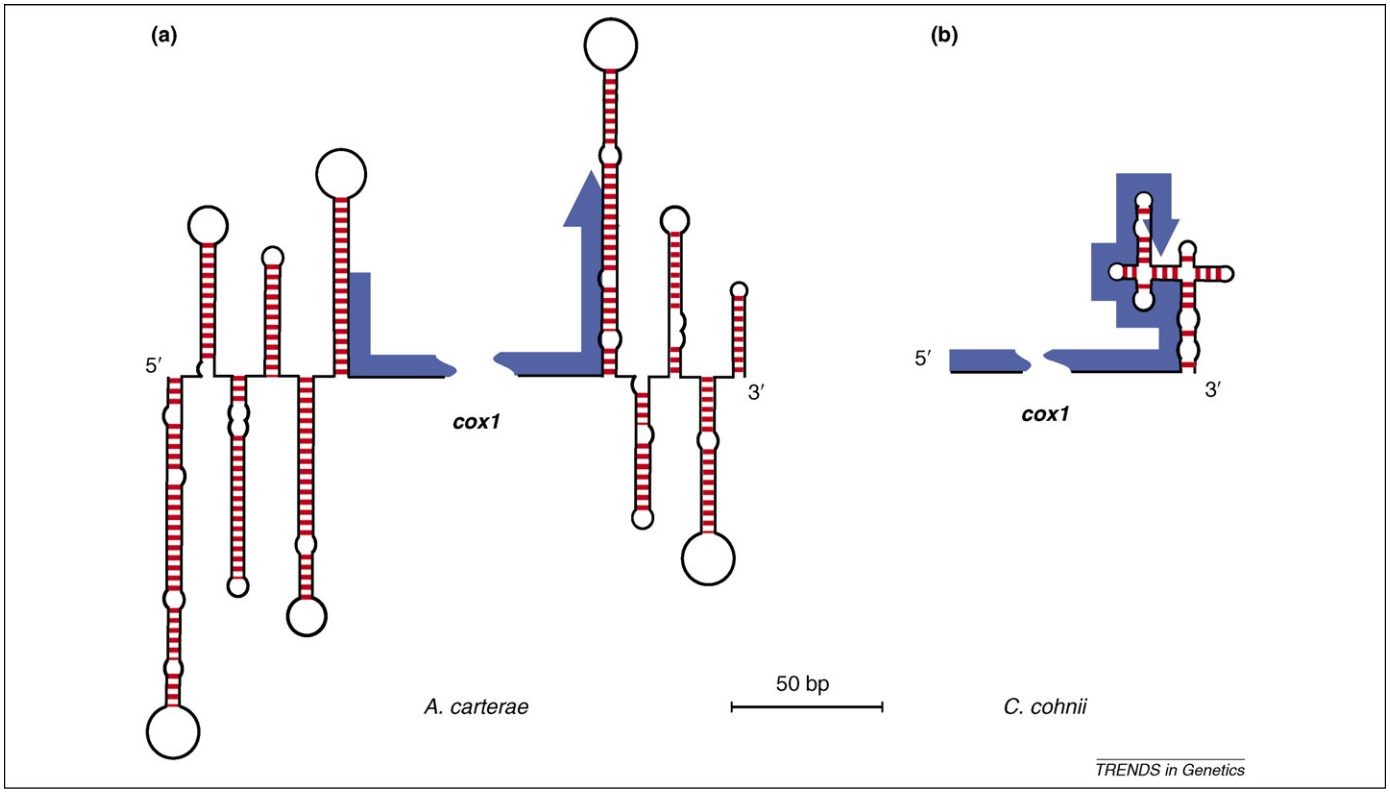

Figure 2. Inverted repeats in dinoflagellate mitochondrial DNA (mtDNA). Inverted repeats have been reported for the 5′ and 3′ ends of the *cox1* gene in **(a)** *Amphidinium carterae* [21] and **(b)** *Crypthecodinium cohnii* [15]. Only one DNA strand is depicted in each case. The resultant *cox1* transcript is shown in blue.

differed from one another (Figure 1). Divergence was marked by a complete loss of significant similarity at defined points rather than a gradual loss of correspondence, suggesting that these different arrangements of flanking sequence were generated by recombination. Recombination-mediated genomic rearrangements are commonplace in the mtDNA of land plants, although this lineage is not closely related to dinoflagellates (reviewed in Ref. [32]).

Attempts to map the genes relative to each other, whether by sequencing of mtDNA fragments or through PCR using primers from different genes, gave apparently conflicting results for most species [16,21,23]. With *A. carterae*, for example, it appeared that *cox3* was adjacent to *cob* in a head-to-head configuration, with a range of different intergenic sequences. Similarly, *cob* appeared to be adjacent to *cox1* in a tail-to-tail configuration. However, no PCR product could be obtained running from *cox3* through *cob* to *cox1*, although such a sequence should be short enough to be amplifiable by PCR (<3 kbp) [21]. Similar analyses of the *O. marina* genome indicated the presence of tandem arrays of individual genes [22] but a similar lack of any three-gene arrangements.

The sequencing studies indicated that a relatively large amount of mtDNA (probably ~85% for *A. carterae*) is noncoding, again reminiscent of the large amount of noncoding sequence in land plants [32]. The most striking aspect of the noncoding DNA in *A. carterae*, *C. cohnii* and *K. micrum* is the extraordinary density of inverted repeat structures, packed so closely together that there are sometimes few or no nucleotides separating one inverted repeat from another. The typical size of inverted repeats varies among dinoflagellate species, from a few base pairs in *C. cohnii* [15] to many tens in *A. carterae* [21]. Inverted repeats can also be formed from coding sequence and their corresponding 5′ or 3′ flanking regions [15,21] (Figure 2), but are not found wholly within genes. Such repeats could allow the formation of stem-loop structures in both DNA strands. In other organisms, stem-loop structures are thought to play roles in the control of mitochondrial replication [33], transcript stability [34] and genome recombination [35], but their role in dinoflagellates remains unknown.

RNA genes

In the mitochondrial genome of Apicomplexa such as *Plasmodium* and *Theileria*, rRNAs are encoded in pieces as a series of short fragments on both DNA strands [26]. However, these fragments comprise only a fraction of the rRNA sequence found in other organisms; how Apicomplexa are able to assemble these sequences to form functional ribosomes is not fully understood. A similar situation is likely to exist in dinoflagellate mitochondria. Short sequences (tens to a couple of hundred nucleotides) with clear similarity to large subunit (LSU) rRNA sequences from Apicomplexa are found scattered throughout the mitochondrial genomes of *A. catenella*, *C. cohnii*, *K. micrum* and *O. marina*. To date, only one SSU (small subunit) rRNA sequence fragment has been identified in *K. micrum* [23]. The failure to identify more is most likely caused by the small size, high divergence and fragmentation of such sequences. For similar reasons, there was uncertainty for some time over whether ribosomal DNA (rDNA) was present in the dinoflagellate chloroplast genome [36]. None

of the studies on dinoflagellate mitochondrial DNA has identified any tRNA genes, indicating that tRNAs must be imported. This process has been reported in many organisms including the apicomplexan *Toxoplasma gondii* [37], certain other protozoa [38] and plants [39]. It has been suggested that apicomplexan mitochondria obtain tRNA-fMet from the chloroplast [40,41]. The same might be true in dinoflagellates, because tRNA-fMet is one of only a handful of tRNAs encoded in the dinoflagellate chloroplast genome [36,42].

Dinoflagellate mitochondrial genome architecture

Hybridization experiments provide further insights into the overall organization of the dinoflagellate mitochondrial genome. Probes for the *cox1*, *cox3* and *cob* coding sequences hybridized to several restriction fragments of different sizes and stoichiometries in Southern blots of mtDNA from *Gonyaulax polyedra* [16], *A. carterae* [21] and *C. cohnii* [23], consistent with the range of different 5′ and 3′ sequences found flanking the gene coding sequences. Although the uncut DNA fragments are in excess of 10–20 kb, at the resolving limit of simple DNA gel electrophoresis, preliminary studies using Southern blot hybridization of *A. carterae* DNA subjected to pulse-field gel electrophoresis indicated the uncut DNA behaves as molecules of ~30 kb (E. Nash, unpublished data). Because the total amount of DNA sequence determined from this organism was greater than this size, it is likely that the genome consists of several nonidentical DNA molecules of ~30 kb.

mtDNAs in other organisms are extremely diverse in structure. Forms of mtDNA observed so far include circular, linear, tandemly repeated linear and maxi/minicircular forms (see Ref. [43] and references therein). The only mitochondrial genome that appears to be even remotely similar in structure to that of dinoflagellates is found in the ichthyosporean *Amoebidium parasiticum*, where genes are present in varying arrangements on a series of linear chromosomes totalling >200 kb [44].

Dinoflagellate mitochondrial transcripts are also unusual

Several EST datasets have been generated from polyadenylated dinoflagellate transcripts. The *cox1*, *cox3* and *cob* sequences are well represented in these datasets, indicating that 3′ polyadenylation occurs with dinoflagellate mitochondrial transcripts, as has been reported for *Plasmodium* sp. [45]. EST data indicate that all *cox1* and *cob* transcripts are polyadenylated at the same position, probably with little or no 3′ untranslated region. In addition, *O. marina* transcripts have a 5′ oligo-U cap [22], although this phenomenon has not yet been reported from the other dinoflagellates. Transcripts containing multiple gene fragments have been observed in *G. polyedra* [16] and *D. baltica* [24].

Two of the recent studies of dinoflagellate mtDNA attempted to amplify a full-length copy of *cox3* from mtDNA (based on the EST data) but failed [21,23]. This failure was explained by the finding that the *K. micrum* *cox3* gene is bipartite, separated into two gene fragments [23]. The full-length *cox3* transcript is generated by *trans*-splicing a 712 nucleotide 5′ fragment with a 119 nucleotide 3′ fragment. The splice site in the *K. micrum* *cox3* transcript contains five adenine nucleotides, which seem to be incorporated from the poly-A tail of the prespliced 5′ transcript [23]. Mature *A. carterae* *cox3* transcripts do not contain any nucleotides derived from the polyA tail at the splice site [21]. *Trans*-spliced transcripts have previously been found for mitochondrial genes in land plants (reviewed in Ref. [46]), the protist *Diplonema* sp. [47], as well as for nuclear genes in dinoflagellates [48].

There is no evidence, based on alignments of predicted protein sequence in regions upstream of conserved sequence, for conventional start codons in the 5′ region of dinoflagellate mitochondrial transcripts. In *trans*-spliced *cox3* transcripts, poly-A addition to a uracil nucleotide at the 3′ end creates a new stop codon—an in-frame UAA—in all species thus far examined (*A. carterae, C. cohnii, K. micrum* and *O. marina*) [21–23]. Generation of a stop codon by polyadenylation has also been observed in mammalian mitochondrial transcripts [49]. However, in *cox1* and *cob* transcripts, the open reading frame is followed by a poly-A tail without a recognizable stop codon. Although dinoflagellate mitochondrial translation might recognize UAA as a stop codon in *cox3* transcripts, this seems not to be necessary in *cox1* or *cob* transcripts. Although the use of nonconventional start codons is observed in ciliate [28,50] and apicomplexan [25,26] mitochondria, both of these lineages use conventional stop codons encoded in the DNA.

Transcript editing, in which transcribed RNA is modified by sequence alteration before undergoing translation (Figure 3), has been inferred for all the dinoflagellate mitochondria studied (except *O. marina*) by comparing cDNAs with genomic sequence. The substitutions observed are ~50% A→G (or A→I) but include many other changes (primarily transitions), mostly at the first and second codon positions; the G→C substitutions seem to be unique to dinoflagellate mitochondrial genes [51]. Comparison of mitochondrial cDNA sequences from EST databases of *A. carterae* and *K. micrum* reveals that some transcripts are partially edited, because not all the editable sites have been altered [21,23]. The distribution of edited sites shows that editing does not proceed systematically in one direction along a transcript. No information concerning the mechanism of RNA editing in dinoflagellates, or the possible enzymes involved, is yet available; however, all substitutions are assumed to occur at the RNA level. Although the possibility remains that edited transcripts arise from alternative gene copies that differ as a result of accumulated mutations, this seems unlikely given the prevalence in editing of A→G substitutions, the preference for changes at first and second codon positions and the conservation of sites between species.

Unusual features of mtDNA-encoded proteins

Cytochrome oxidase (mitochondrial complex IV) consists of three major polypeptide subunits, designated Cox1, 2 and 3. The complex uses energy from electrons passed to it by cytochrome *c* to pump protons across the inner mitochondrial membrane. Although Cox3 is not directly involved in proton pumping, it interacts with Cox1 at its N terminus [52]. The predicted amino acid sequences of mitochond-

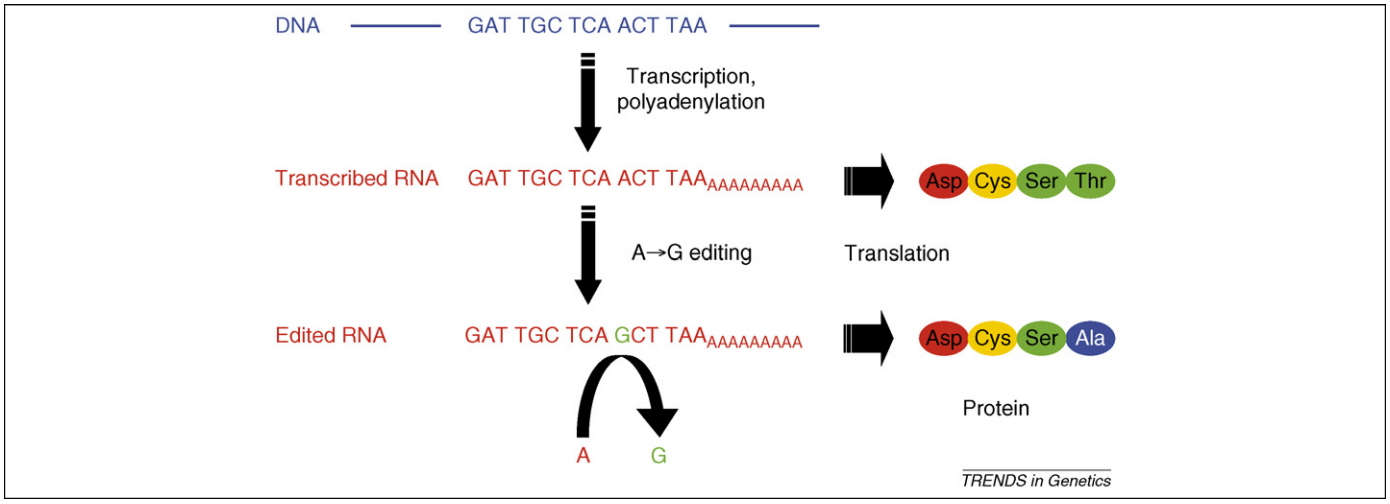

Figure 3. Transcript editing. An example of how RNA editing might result in a change in identity of the amino acid encoded by the genome. Here, an A→G substitution in a first codon position changes the amino acid encoded in the transcript from alanine to threonine. To emphasize the residue involved in editing, T nucleotides in the DNA are also shown as T in the RNA.

rially encoded dinoflagellate proteins show substitutions at several functionally important sites that are conserved in most other organisms. These include residues positioned at the Cox1–Cox3 interface within the cytochrome oxidase complex that are known to have roles in proton pumping across the inner mitochondrial membrane. The consequences of these substitutions for protein function remain unclear.

The most striking feature of protein sequences predicted from the mtDNA data is found in *O. marina*, where the *cob* and *cox3* coding regions are fused both in transcripts and in genomic DNA [22]. Both coding regions appear to be full length, in frame, and without intervening stop codons. Assuming that this represents the expressed *cob* and *cox3* sequences, is such a fusion protein functionally possible? The C terminus of Cob and the N terminus of Cox3 both reside on the matrix side of the inner mitochondrial membrane, so the ~30 residues separating Cox3 and Cob regions of the fused protein might be sufficient to link the two subunits. However, the fusion of these proteins might impose steric restrictions on their function or their assembly into functional complexes in the inner mitochondrial membrane, and the covalent linkage of two electron transport chain complexes would be, to our knowledge, unprecedented.

Alternatively, the *O. marina* Cob and Cox3 proteins might be cleaved after synthesis. There are possible precedents for this in fusions between Cox subunits in other organisms. A translational fusion of *cox1* and *cox2* is found in the soil amoeba *Acanthamoeba castellanii*. Analysis of the mature Cox2 protein in *A. castellanii* suggests that it is not fused to Cox1 [53], and it is likely that the fusion protein is cleaved following synthesis. Outside of the eukaryotes, the hyperthermophilic bacterium *Thermus thermophilus* contains a Cox1–Cox3 fusion protein, but this protein might also undergo posttranslational cleavage [54].

Concluding remarks

The overall picture is of a very unusual genome organization in dinoflagellate mitochondria. The mitochondrial genome is probably subject to a high level of intramolecular recombination, perhaps reflecting the large amount of inverted repeat sequence in the noncoding regions. This is somewhat reminiscent of the situation in land plants, in which recombination often fragments a large mitochondrial genome into small subgenomic circles [32]. Recombination in the dinoflagellate mitochondrion generates a large number of rearranged gene fragments with a range of different flanking sequences floating like driftwood in a sea of noncoding material. This complex mass of material is distributed over a range of molecules. For the cell to make sense of this molecular flotsam requires a large amount of posttranscriptional modification, including editing and *trans*-splicing.

How did all this evolve [14–24]? The small number of genes in dinoflagellate and apicomplexan mtDNA compared to the larger number in ciliates suggests that the ancestor of dinoflagellates and Apicomplexa transferred the large majority of its mitochondrial genes to the nucleus after diverging from the ciliates and before splitting into the separate dinoflagellate and apicomplexan lineages. Mitochondrial genome reorganisation occurred in dinoflagellates after divergence from Apicomplexa [21–23]. Among the dinoflagellates, the organization of mtDNA seems to have degenerated less in *O. marina* than the others. Cox3 is encoded in one piece in *O. marina*, there is no transcript editing and the arrangement of genes seems to be less scrambled than in other dinoflagellate species [22]. This might be consistent with studies using nuclear sequences that place *O. marina* as a basal lineage within the dinoflagellate group [31]. However, why mitochondrial genome deterioration should have become frozen in *O. marina* after the divergence of this lineage is not clear. Given the variations between mtDNAs of different dinoflagellate species thus far described, greater taxon sampling is needed. It will be particularly interesting to examine the mtDNA of *Perkinsus marinus,* thought to be the most basal dinoflagellate lineage thus far discovered [31], especially because a nuclear genome project for this organism is now underway.

Together, these studies give a detailed understanding of genome disorganisation in dinoflagellate mitochondria.

Box 2. Outstanding questions

Why have so many mitochondrial genes been transferred to the nucleus in dinoflagellates and Apicomplexa?
How is the dinoflagellate mitochondrial genome organized? Is it circular or linear, and how are the genes and pseudogenes organized within it?
Why do dinoflagellates contain such high levels of noncoding mitochondrial DNA (mtDNA) compared with apicomplexan mitochondria and why is dinoflagellate gene arrangement so much more chaotic?
What is the function, if any, of the mtDNA inverted repeats?
What is the function of the oligo-U cap that is attached to the 5′ end of mitochondrial transcripts in *Oxyrrhis marina*?
How are *trans*-splicing and RNA editing directed?
How are ribosomes assembled using the fragmentary ribosomal RNAs?
What codons are used for translation initiation and termination? How does termination take place, if conventional stop codons are not used? Are transfer RNAs imported from the chloroplast or the cytosol?
Do dinoflagellates possess a functioning Cob–Cox3 fusion protein? How might the fusion affect electron transport?
What does the mitochondrial genome of *Perkinsus marinus* (an early diverging dinoflagellate species) look like?

The unusual nature of the dinoflagellate mitochondrial genome compares with the atypical features previously discovered for the nuclear and chloroplast genomes. Although the structure of the mitochondrial genome is now fairly clear, many questions remain (Box 2). Novel features in mtDNA often appear from studies based on a single species (e.g. Ref. [8]). In dinoflagellates, the recent studies outlined here demonstrate that these characteristics exist in a range of species and that most are common to all species.

Acknowledgements

Work in the authors' laboratory is supported by the Biotechnology and Biological Sciences Research Council, the Leverhulme Trust and the Newton Trust.

References

1 Embley, T.M. and Martin, W. (2006) Eukaryotic evolution, changes and challenges. *Nature* 440, 623–630
2 Martin, W. and Muller, M. (1998) The hydrogen hypothesis for the first eukaryote. *Nature* 392, 37–41
3 Tielens, A.G. *et al.* (2002) Mitochondria as we don't know them. *Trends Biochem. Sci.* 27, 564–572
4 Embley, T.M. *et al.* (2003) Mitochondria and hydrogenosomes are two forms of the same fundamental organelle. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 191–202
5 Boxma, B. *et al.* (2005) An anaerobic mitochondrion that produces hydrogen. *Nature* 434, 74–79
6 Stechmann, A. *et al.* (2008) Organelles in *Blastocystis* that blur the distinction between mitochondria and hydrogenosomes. *Curr. Biol.* 18, 580–585
7 Ward, B.L. *et al.* (1981) The mitochondrial genome is large and variable in a family of plants (cucurbitaceae). *Cell* 25, 793–803
8 Lang, B.F. *et al.* (1997) An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* 387, 493–497
9 Unseld, M. *et al.* (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat. Genet.* 15, 57–61
10 Gray, M.W. *et al.* (2004) Mitochondria of protists. *Annu. Rev. Genet.* 38, 477–524
11 Taylor, F.J.R., ed. (1987) *The Biology of Dinoflagellates*, Blackwell Scientific
12 Baker, A.C. (2003) Flexibility and specificity in coral-algal symbiosis: diversity, ecology, and biogeography of *Symbiodinium*. *Annu. Rev. Ecol. Evol. Syst.* 34, 661–689
13 Kirkpatrick, B. *et al.* (2004) Literature review of Florida red tide: implications for human health effects. *Harmful Algae* 3, 99–115
14 Norman, J.E. and Gray, M.W. (1997) The cytochrome oxidase subunit 1 gene (*cox1*) from the dinoflagellate, *Crypthecodinium cohnii*. *FEBS Lett.* 413, 333–338
15 Norman, J.E. and Gray, M.W. (2001) A complex organization of the gene encoding cytochrome oxidase subunit 1 in the mitochondrial genome of the dinoflagellate, *Crypthecodinium cohnii*: homologous recombination generates two different *cox1* open reading frames. *J. Mol. Evol.* 53, 351–363
16 Chaput, H. *et al.* (2002) Polyadenylated transcripts containing random gene fragments are expressed in dinoflagellate mitochondria. *Protist* 153, 111–122
17 Zhang, H. and Lin, S. (2002) Detection and quantification of *Pfiesteria piscicida* by using the mitochondrial cytochrome *b* gene. *Appl. Environ. Microbiol.* 68, 989–994
18 Zhang, H. and Lin, S. (2005) Development of a *cob*-18S rRNA gene real-time PCR assay for quantifying *Pfiesteria shumwayae* in the natural environment. *Appl. Environ. Microbiol.* 71, 7053–7063
19 Zhang, H. and Lin, S. (2005) Mitochondrial cytochrome *b* mRNA editing in dinoflagellates: possible ecological and evolutionary associations? *J. Eukaryot. Microbiol.* 52, 538–545
20 Kamikawa, R. *et al.* (2007) Fragmentation of mitochondrial large subunit rRNA in the dinoflagellate *Alexandrium catenella* and the evolution of rRNA structure in alveolate mitochondria. *Protist* 158, 239–245
21 Nash, E.A. *et al.* (2007) Organization of the mitochondrial genome in the dinoflagellate *Amphidinium carterae*. *Mol. Biol. Evol.* 24, 1528–1536
22 Slamovits, C.H. *et al.* (2007) The highly reduced and fragmented mitochondrial genome of the early-branching dinoflagellate *Oxyrrhis marina* shares characteristics with both apicomplexan and dinoflagellate mitochondrial genomes. *J. Mol. Biol.* 372, 356–368
23 Jackson, C.J. *et al.* (2007) Broad genomic and transcriptional analysis reveals a highly derived genome in dinoflagellate mitochondria. *BMC Biol.* 5, 41
24 Imanian, B. and Keeling, P.J. (2007) The dinoflagellates *Durinskia baltica* and *Kryptoperidinium foliaceum* retain functionally overlapping mitochondria from two evolutionarily distinct lineages. *BMC Evol. Biol.* 7, 172
25 Feagin, J.E. (1992) The 6 kb element of *Plasmodium falciparum* encodes mitochondrial cytochrome genes. *Mol. Biochem. Parasitol.* 52, 145–148
26 Kairo, A. *et al.* (1994) A 7.1 kb linear DNA molecule of Theileria parva has scrambled rDNA sequences and open reading frames for mitochondrially encoded proteins. *EMBO J.* 13, 898–905
27 Wilson, R.J. and Williamson, D.H. (1997) Extrachromosomal DNA in the Apicomplexa. *Microbiol. Mol. Biol. Rev.* 61, 1–16
28 Burger, G. *et al.* (2000) Complete sequence of the mitochondrial genome of *Tetrahymena pyriformis* and comparison with *Paramecium aurelia* mitochondrial DNA. *J. Mol. Biol.* 297, 365–380
29 Ziaie, Z. and Suyama, Y. (1987) The cytochrome oxidase subunit I gene of *Tetrahymena*: a 57 amino acid NH_2-terminal extension and a 108 amino acid insert. *Curr. Genet.* 12, 357–368
30 Leander, B.S. and Keeling, P.J. (2004) Early evolutionary history of dinoflagellates and apicomplexans (Alveolata) as inferred from hsp90 and actin phylogenies. *J. Phycol.* 40, 341–350
31 Saldarriaga, J.F. *et al.* (2003) Multiple protein phylogenies show that *Oxyrrhis marina* and *Perkinsus marinus* are early branches of the dinoflagellate lineage. *Int. J. Syst. Evol. Microbiol.* 53, 355–365
32 Knoop, V. (2004) The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Curr. Genet.* 46, 123–139
33 Arunkumar, K.P. and Nagaraju, J. (2006) Unusually long palindromes are abundant in mitochondrial control regions of insects and nematodes. *PLoS ONE* 1, e110
34 Kuhn, J. *et al.* (2001) Transcript lifetime is balanced between stabilizing stem-loop structures and degradation-promoting polyadenylation in plant mitochondria. *Mol. Cell. Biol.* 21, 731–742
35 Bartoszewski, G. *et al.* (2004) Organization of repetitive DNAs and the genomic regions carrying ribosomal RNA, *cob*, and *atp9* genes in the cucurbit mitochondrial genomes. *Theor. Appl. Genet.* 108, 982–992
36 Barbrook, A.C. *et al.* (2006) Comparative analysis of dinoflagellate chloroplast genomes reveals rRNA and tRNA genes. *BMC Genomics* 7, 297

37 Esseiva, A.C. *et al.* (2004) Mitochondrial tRNA import in *Toxoplasma gondii*. *J. Biol. Chem.* 279, 42363–42368
38 Goswami, S. *et al.* (2006) A bifunctional tRNA import receptor from *Leishmania* mitochondria. *Proc. Natl. Acad. Sci. U. S. A.* 103, 8354–8359
39 Delage, L. *et al.* (2003) *In vitro* import of a nuclearly encoded tRNA into mitochondria of *Solanum tuberosum*. *Mol. Cell. Biol.* 23, 4000–4012
40 Barbrook, A.C. *et al.* (2006) Why are plastid genomes retained in non-photosynthetic organisms? *Trends Plant Sci.* 11, 101–108
41 Howe, C.J. and Purton, S. (2007) The little genome of apicomplexan plastids: its *raison d'etre* and a possible explanation for the 'delayed death' phenomenon. *Protist* 158, 121–133
42 Nelson, M.J. *et al.* (2007) Identification and transcription of transfer RNA genes in dinoflagellate plastid minicircles. *Gene* 392, 291–298
43 Burger, G. *et al.* (2003) Mitochondrial genomes: anything goes. *Trends Genet.* 19, 709–716
44 Burger, G. *et al.* (2003) Unique mitochondrial genome architecture in unicellular relatives of animals. *Proc. Natl. Acad. Sci. U. S. A.* 100, 892–897
45 Gillespie, D.E. *et al.* (1999) The fragmented mitochondrial ribosomal RNAs of *Plasmodium falciparum* have short A tails. *Nucleic Acids Res.* 27, 2416–2422
46 Bonen, L. (2008) *cis*- and *trans*-splicing of group II introns in plant mitochondria. *Mitochondrion* 8, 26–34
47 Marande, W. and Burger, G. (2007) Mitochondrial DNA as a genomic jigsaw puzzle. *Science* 318, 415
48 Zhang, H. *et al.* (2007) Spliced leader RNA *trans*-splicing in dinoflagellates. *Proc. Natl. Acad. Sci. U. S. A.* 104, 4618–4623
49 Anderson, S. *et al.* (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290, 457–465
50 Edqvist, J. *et al.* (2000) Expression of mitochondrial protein-coding genes in *Tetrahymena pyriformis*. *J. Mol. Biol.* 297, 381–393
51 Gray, M.W. (2003) Diversity and evolution of mitochondrial RNA editing systems. *IUBMB Life* 55, 227–233
52 Hosler, J.P. (2004) The influence of subunit III of cytochrome *c* oxidase on the D pathway, the proton exit pathway and mechanism-based inactivation in subunit I. *Biochim. Biophys. Acta* 1655, 332–339
53 Lonergan, K.M. and Gray, M.W. (1996) Expression of a continuous open reading frame encoding subunits 1 and 2 of cytochrome *c* oxidase in the mitochondrial DNA of *Acanthamoeba castellanii*. *J. Mol. Biol.* 257, 1019–1030
54 Mather, M.W. *et al.* (1993) Cytochrome oxidase genes from *Thermus thermophilus*. Nucleotide sequence of the fused gene and analysis of the deduced primary structures for subunits I and III of cytochrome caa_3. *J. Biol. Chem.* 268, 5395–5408
55 Rizzo, P.J. (2003) Those amazing dinoflagellate chromosomes. *Cell Res.* 13, 215–217
56 Wong, J.T.Y. *et al.* (2003) Histone-like proteins of the dinoflagellate *Crypthecodinium cohnii* have homologies to bacterial DNA-binding proteins. *Eukaryot. Cell* 2, 646–650
57 Loeblich, A.R., 3rd (1984) Dinoflagellate evolution. In *Dinoflagellates* (Spector, D.L., ed.), pp. 481–522, Academic Press
58 Fast, N.M. *et al.* (2002) Re-examining alveolate evolution using multiple protein molecular phylogenies. *J. Eukaryot. Microbiol.* 49, 30–37
59 Boczar, B.A. *et al.* (1991) Characterization of satellite DNA from three marine dinoflagellates (Dinophyceae): *Glenodinium* sp. and two members of the toxic genus. *Plant Physiol.* 97, 613–618
60 Barbrook, A.C. and Howe, C.J. (2000) Minicircular plastid DNA in the dinoflagellate *Amphidinium operculatum*. *Mol. Gen. Genet.* 263, 152–158
61 Zhang, Z. *et al.* (1999) Single gene circles in dinoflagellate chloroplast genomes. *Nature* 400, 155–159
62 Nisbet, R.E.R. *et al.* (2008) Transcript analysis of dinoflagellate plastid gene minicircles. *Protist* 159, 31–39
63 Bhaud, Y. *et al.* (2000) Morphology and behaviour of dinoflagellate chromosomes during the cell cycle and mitosis. *J. Cell Sci.* 113, 1231–1239
64 Zhang, H. *et al.* (2007) A three-gene dinoflagellate phylogeny suggests monophyly of prorocentrales and a basal position for *Amphidinium* and *Heterocapsa*. *J. Mol. Evol.* 65, 463–474

BSHG
The British Society for Human Genetics

British Human Genetics Conference
15–17 September 2008, University of York, UK

Scientific Programme
Developmental Genetics, Genetic Susceptibility Testing,
Mechanisms of Disease, Genomic Disorders,
Aneuploidy, Ciliopathies and Prenatal Diagnosis

Further information
The Conference Office, British Society for Human Genetics,
Clinical Genetics Unit, Birmingham Women's Hospital, Edgbaston,
Birmingham, B15 2TG, UK
Tel: +44 (0)121 627 2634 Fax: +44 (0)121 623 6971
Email: york2008@bshg.org.uk Website: www.bshg.org.uk