Dinoflagellate Genome Evolution

Jennifer H. Wisecaver and Jeremiah D. Hackett

Ecology and Evolutionary Biology Department, University of Arizona, Tucson, Arizona 85721; email: hackettj@email.arizona.edu

Annu. Rev. Microbiol. 2011. 65:369–87

First published online as a Review in Advance on June 14, 2011

The *Annual Review of Microbiology* is online at micro.annualreviews.org

This article's doi: 10.1146/annurev-micro-090110-102841

0066-4227/11/1013-0369$20.00

Keywords

endosymbiosis, *trans*-splicing, gene transfer, alveolate, gene expression

Abstract

The dinoflagellates are an ecologically important group of microbial eukaryotes that have evolved many novel genomic characteristics. They possess some of the largest nuclear genomes among eukaryotes arranged on permanently condensed liquid-crystalline chromosomes. Recent advances have revealed the presence of genes arranged in tandem arrays, *trans*-splicing of messenger RNAs, and a reduced role for transcriptional regulation compared to other eukaryotes. In contrast, the mitochondrial and plastid genomes have the smallest gene content among functional eukaryotic organelles. Dinoflagellate biology and genome evolution have been dramatically influenced by lateral transfer of individual genes and large-scale transfer of genes through endosymbiosis. Next-generation sequencing technologies have only recently made genome-scale analyses of these organisms possible, and these new methods are helping researchers better understand the biology and evolution of this enigmatic group of eukaryotes.

Contents

INTRODUCTION

Dinoflagellates are ecologically and economically important organisms in aquatic environments. They display tremendous morphological diversity and have one of the most extensive fossil records among microbial eukaryotes due to the formation of robust cysts by a large number of species (33). In marine environments they are important primary producers, both as free-living phytoplankton and as symbionts of reef-forming corals. Many species are heterotrophic or mixotrophic and are prolific grazers of planktonic organisms. Dinoflagellates also produce a wide variety of secondary metabolites including a diverse array of toxins that have a significant impact on marine ecosystems and fisheries.

Whereas some aspects of eukaryotes' (particularly microbial lineages) genomes differ from canonical eukaryotic traits, the dinoflagellates are notable for the extreme nature and sheer number of novel genome characteristics. They are perhaps the most ancient and diverse group of eukaryotes for which there is not yet a complete genome sequence. Next-generation sequencing technologies have finally placed a complete genome sequence of a dinoflagellate within reach and are already helping to elucidate the biology and evolution of these enigmatic organisms.

Plastid: organelles of eukaryotes involved in photosynthesis

DINOFLAGELLATE PHYLOGENETICS

Dinoflagellates, apicomplexans, and ciliates are the dominant phyla of the well-supported superphylum Alveolata, named for the flattened vesicles (cortical alveoli) that form a continuous layer just under the plasma membrane (21). The ciliates are predominantly unicellular heterotrophs that exhibit nuclear dimorphism and show deviations from the universal genetic code (87, 90, 113). Apicomplexans are mostly intracellular parasites (54) and contain a nonphotosynthetic plastid (the apicoplast) involved in production of fatty acids, isoprene, heme, and iron-sulfur clusters (119). Phylogenetic analyses are inconclusive but suggest alveolates are related to stramenopiles (e.g., diatoms and kelp) and rhizarians (e.g., forams and radiolarians) (18, 19). Within the alveolates, dinoflagellates are sister to the apicomplexans (29). Alveolate species that do not fall within these three phyla are important for inferring ancestral conditions. For example, the oyster pathogen *Perkinsus marinus* is sister to dinoflagellates and retains many typical eukaryotic characteristics (91), whereas the free-living marine phototroph *Chromera velia* is sister to the parasitic apicomplexans (70). Molecular clock analyses indicate the dinoflagellates and apicomplexans diverged 800–900 million years ago (11, 37), consistent with fossils attributed to dinoflagellates appearing as early as the late Mesoproterozoic (65).

The dinoflagellates can be divided into three main groups based on molecular phylogenies: Oxyrrhinales, Syndiniales, and the core dinoflagellates comprising the majority of characterized species (67) (**Figure 1**). A large unculturable diversity of marine dinoflagellates, termed Marine Alveolates Group I (MAGI),

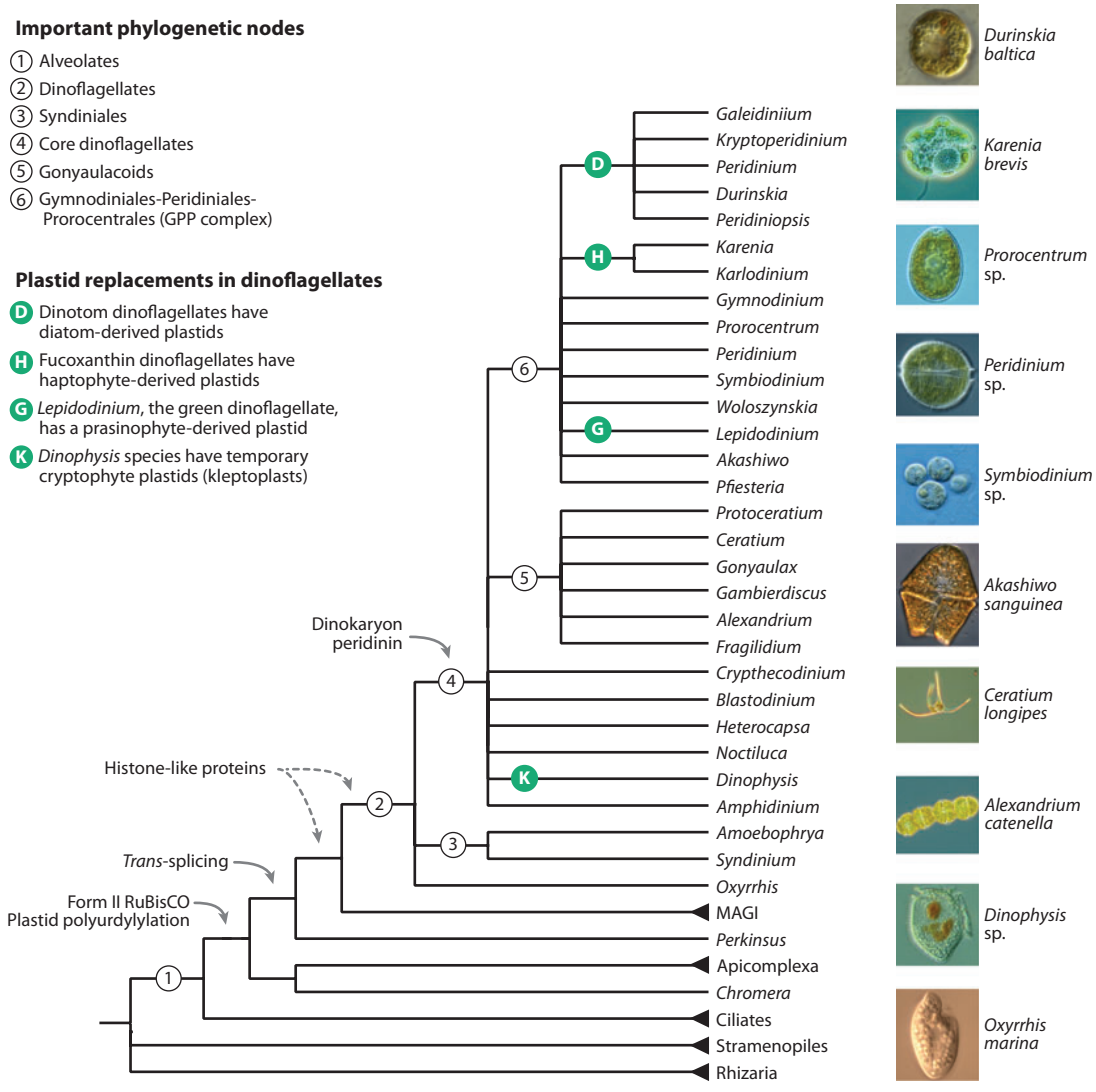

Figure 1

A phylogenetic model of dinoflagellate evolutionary relationships based on molecular data. Branches uniting supported clades are indicated with numbers. Plastid replacements are indicated with letters. Arrows indicate the most likely branch for the origin of some distinctive characteristics of dinoflagellates. Micrographs were obtained from the micro*scope Web site (**http://starcentral.mbl.edu/microscope/portal.php**) and used with permission. Image copyrights: *Durinskia baltica* (S. Murray, M. Hoppenrath, J. Larsen & D. Patterson); *Karenia brevis*, *Prorocentrum* sp., *Ceratium longipes*, *Alexandrium catenella*, and *Dinophysis* sp. (R. Andersen & D. Patterson); *Peridinium* sp. (M. Bahr & D. Patterson); *Symbiodinium* sp. (D. Patterson & M. Farmer); *Akashiwo sanguinea* (H. Su Yoon); *Oxyrrhis marina* (R. Moore & M.V. Sanchez-Puerta). Abbreviation: MAGI, Marine Alveolates Group I.

has been described by 18S rDNA environmental surveys and appears to be composed of primarily parasitic species (60, 69, 105). Syndiniales, such as *Amoebophrya*, are also predominantly parasitic species. The free-living and heterotrophic Oxyrrhinales contain just one morphospecies, *Oxyrrhis marina* (61). These heterotrophic taxa (MAGI, the Syndiniales, and *Oxyrrhis*) occupy a basal position in the dinoflagellate tree (41, 95, 97, 124). However,

phylogenetic analyses have not determined the positions of these groups relative to each other, leaving the deepest branches of the dinoflagellate tree unresolved. This is problematic when attempting to infer ancestral states of some of the more unusual dinoflagellate characteristics discussed below.

Within the core dinoflagellates, the traditional orders were based on morphological characters such as thecal plate tabulation patterns or the absence of theca altogether (31). Molecular phylogenetics has shown that many of these orders are poly- and paraphyletic and that thecal characteristics have evolved multiple times within the core dinoflagellates (67, 124). Resolving the internal relationships among dinoflagellates has been difficult. **Figure 1** shows the few supported clades within the core dinoflagellates, including monophyly of the gonyaulacoids, the Gymnodiniales-Peridiniales-Prorocentrales complex taxa, and groups with more recently acquired plastids such as the dinotom and fucoxanthin dinoflagellates. The parasitic Blastodiniales and free-living, heterotrophic Noctilucales are particularly unstable in phylogenies, appearing as basal lineages in some gene trees and highly derived in others (35, 41, 97, 99).

NUCLEAR BIOLOGY

Dinoflagellate nuclear structure and biology are among the most divergent in the eukaryotic domain, leading Dodge (26) to consider them mesokaryotes, an intermediate between prokaryotes and eukaryotes, until molecular phylogenetics confirmed their membership in the Alveolata. The primary reason for this designation is the distinctive chromosomes that remain permanently condensed throughout the cell cycle without the aid of nucleosomes. Dinoflagellate genomes are among the largest of any organism and contain the unusual nucleobase hydroxymethyluracil. Recent advances in dinoflagellate genome research include the discoveries of tandem gene arrays, *trans*-spliced mRNAs, and a reduced role of transcriptional regulation compared to other eukaryotes.

Nuclear Genome Organization and Evolution

Most dinoflagellates have dinokaryotic nuclei, distinguished by permanently condensed chromosomes that are attached to the nuclear envelope and lack nucleosomes (12). Amazingly, the chromosome structure appears to be the result of DNA liquid crystal formation in the nucleus (15, 32, 59) (**Figure 2*a***). The nuclear DNA is associated with basic nuclear proteins, or histone-like proteins (HLPs). These proteins have a secondary structure similar to that of bacterial HLPs; however, dinoflagellate HLPs cannot complement the bacterial proteins in vivo (22). Dinoflagellate HLPs are associated with the nucleolus and loops of DNA that extend from the condensed liquid-crystalline chromosomes (22, 94). The interior of the chromosomes comprising the majority of DNA in the nucleus is likely too dense to allow transcription, which is thought to occur instead from genes located on the extrachromosomal loops (103). HLPs affect the compaction of DNA in a concentration-dependent manner, suggesting that they play some role in regulating the condensation of these loops of DNA and the accessibility of genes to transcription factors (22, 94) (**Figure 2*a***).

All core dinoflagellates appear to have this dinokaryotic nuclear structure during some or all stages of their life cycle. Outside of this clade, the presence of these characters is less clear. *Oxyrrhis* appears to have HLPs and lack histones, but it does not have permanently condensed chromosomes. The structure of the nucleus in the Syndiniales has not been thoroughly investigated. Positive chemical tests for basic nuclear proteins in these organisms have been interpreted as evidence of either the presence of histones (20) or HLPs (92).

Dinoflagellates were thought to have completely lost histones; however, recent transcriptome analyses have identified transcripts for all core histones (H2A, H2B, H3, H4) as well as several variants (H2A.X and H2A.Z) (36, 57, 80). These proteins are clearly not involved in packaging the majority of the nuclear DNA,

Dinotoms: dinoflagllates with diatom endosymbionts plastids, which are minimally reduced and retain the diatom nucleus and other organelles

Theca: armored plates composed of cellulose or other polysaccharides formed within the cortical alveoli of dinoflagellates

Dinokaryon: nuclear organization specific to core dinoflagellates characterized by permanently condensed chromosomes that lack nucleosomes

Histone-like protein (HLP): basic nuclear protein found associated with the nuclear DNA of dinoflagellates

Transcriptome: the complete set of RNAs transcribed from the genome of an organism

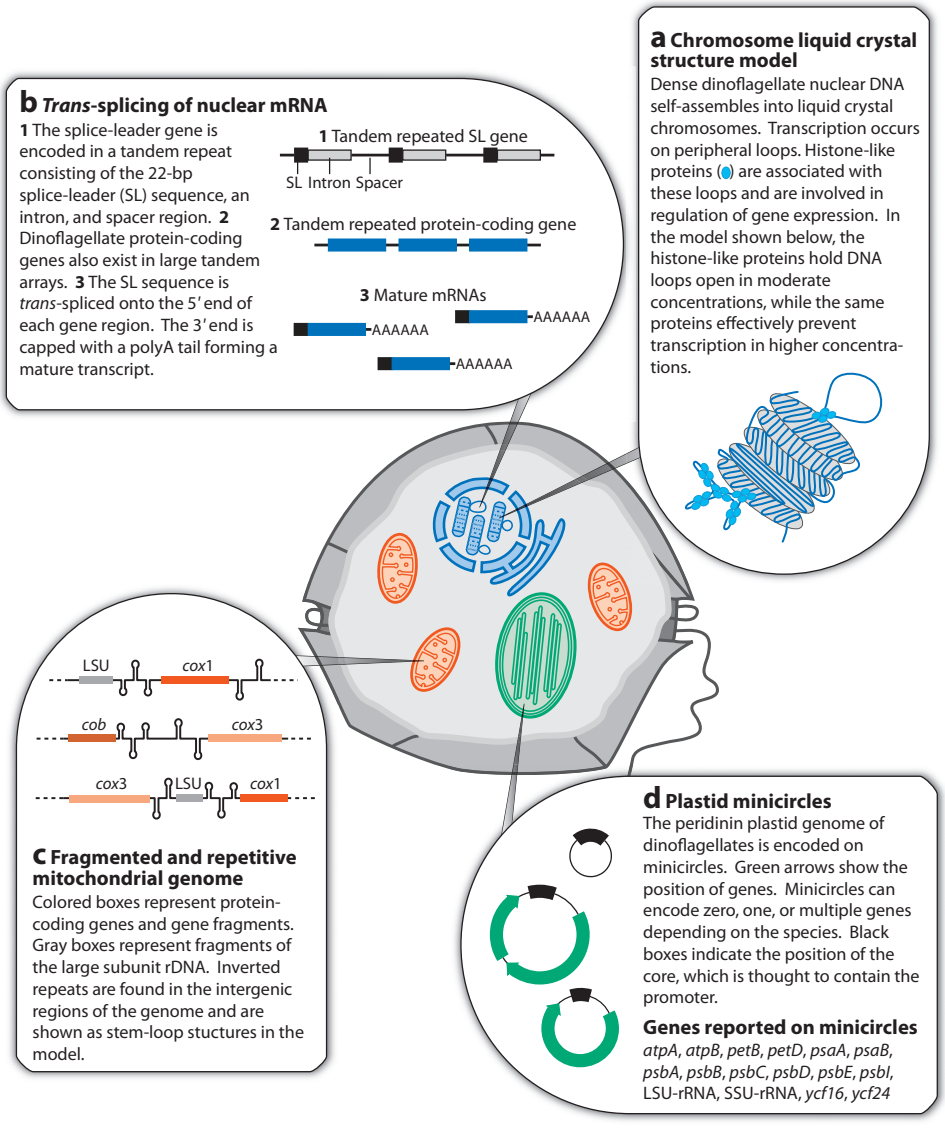

Figure 2

Distinctive elements of dinoflagellate nuclear, mitochondrial, and plastid genomes. (*a*) A model of chromosome structure and regulation of extrachromosomal loop compaction. (*b*) *Trans*-splicing of nuclear mRNAs. (*c*) Model of mitochondrial gene organization. (*d*) Minicircle chromosomes in peridinin-containing dinoflagellates. Abbreviations: LSU, large subunit; SL, splice leader; SSU, small subunit.

at least in the vegetative stage of the life cycle of core dinoflagellates. These histones may be involved in packaging a small portion of the genome and have, thus far, eluded detection or package DNA only in a particular life cycle stage (e.g., cysts).

Dinoflagellates have some of the largest eukaryotic genomes, ranging from 1.5 Gb in species of the coral symbiont *Symbiodinium* to 185 Gb in *Lingulodinium polyedrum* (51, 107). Extreme polyploidy, or the proliferation of repetitive DNA, does not appear to be

Polycistronic transcript: a single molecule of mRNA that contains the information for the translation of several proteins

***Trans*-spliced leader:** a noncoding RNA added to the 5′ end of mature transcripts during mRNA processing

responsible for the large size of these genomes, since DNA reassociation kinetics are similar to that in other eukaryotes (2, 25). A recent genome survey analyzed 230 kb of the *Heterocapsa triquetra* genome, estimated to be 18 to 23 Gb, and found 89.5% of the sequence was nonrepetitive, presumably noncoding DNA (64). Chromosome number varies considerably among species, ranging from 24 to 220 chromosomes that appear to be identical in size and shape within a species (104). In contrast, a karyotype of the predinoflagellate *Perkinsus atlanticus* showed an arrangement of just nine chromosomes (110). A karyotype of a dinoflagellate was attempted in *Pyrocystis lunula* by reconstructing a three-dimensional representation of the nucleus from electron micrographs of 300 serial sections (101). This study revealed 49 pairs of chromosomes with distinctive size and morphology, suggesting diploidy. These results were surprising because dinoflagellates are generally considered to have haploid vegetative cells (31); however, the authors could not rule out genome replication without cell division in the cell cultures.

The complete sequence of a massive nuclear genome from a dinoflagellate has long been out of reach because of the high cost of sequencing using traditional methods. The decrease in cost using next-generation methods now makes it possible to determine the genome sequence of a dinoflagellate. The smallest are near the size of the human genome, which currently costs about $50,000 to sequence, and is rapidly decreasing (13). The first dinoflagellate genome sequence is likely to be completed in the near future and will help answer many of the questions raised in this review.

Another distinctive feature of dinoflagellate genomes is a large portion (over 35%) of the thymine has been replaced by hydroxymethyluracil in many species (25, 89). This base is a natural component of some phage genomes, such as *Bacillus subtilis* phages, but in most eukaryotes, it is the result of oxidative damage of thymine or 5-methylcytosine and is quickly repaired by a DNA glycosylase (14, 40). The role of hydroxymethyluracil in dinoflagellate nuclear biology remains unknown.

Transcription and regulation of gene expression. Several important discoveries have recently been made regarding dinoflagellate transcription. One of the most critical findings was the discovery of 5′ *trans*-splicing of dinoflagellate transcripts. Expressed sequence tags (ESTs) from several dinoflagellates revealed the presence of the same 22-nt leader sequence on the 5′ end of all transcripts produced from the nucleus (55, 125). So far, *trans*-splicing has been found in all species examined, including those outside the core dinoflagellates (i.e., *Oxyrrhis* and *Amoebophrya*) and in *Perkinsus*, indicating that this processing arose early in dinoflagellate evolution (7, 47, 127) (**Figure 1**). The same 22-nt sequence is found on all nuclear mRNAs and is conserved across all species, making it a useful tool for isolating full-length cDNAs. The leader sequences themselves are encoded in tandem gene arrays of unknown number or size (**Figure 2*b***). *Trans*-splicing of an mRNA leader is found for a limited number of mRNAs in several other eukaryotes but is used extensively and is well studied in the trypanosomes (108). Though the extensive use of 5′ *trans*-splicing in dinoflagellates and trypanosomes is likely the result of convergent evolution, the well-studied biology of trypanosomes provides a model for understanding dinoflagellate genome organization and regulation of gene expression. In trypanosomes, genes are organized into clusters and cotranscribed into polycistronic RNAs that are processed into individual mRNAs by *trans*-splicing and polyadenylation. Individual genes lack promoters and most regulation of gene expression is posttranscriptional. The 5′ *trans*-spliced leader and other proteins that interact with motifs in the 3′-untranslated region are responsible for regulating mRNA stability and their translation into proteins (for review see Reference 82). Ongoing and future research in this area is focused on testing

for similarities between trypanosome and dinoflagellate molecular biology.

Several other aspects of dinoflagellate nuclear genomes are similar to those of trypanosomes. Genes are often encoded in tandem arrays encoding virtually identical proteins (6, 58). Bachvaroff & Place (6) determined the genomic structure of 47 genes, identified by EST sequencing, that varied in their expression level as inferred from their frequency in the cDNA library. They found that more highly expressed genes tended to be found in large tandem gene arrays, whereas genes expressed at a lower level appeared to be encoded by only a single gene and to contain more introns. They also identified a putative polyadenylation signal (AAAAG/C) occurring at the site of polyadenylation. Gene amplification may be an important process in the regulation of gene expression over evolutionary timescales in dinoflagellates. If the majority of gene expression is regulated posttranscriptionally in dinoflagellates as in trypanosomes, amplification of genes on the genome could be a primary mechanism for regulating transcript abundance in the cell. Individual genes in these arrays also appear to lack RNA polymerase II promoters. No upstream TATA-box promoter region has been identified near dinoflagellate genes; however, a divergent TATA binding protein (TBP) was found in *Crypthecodinium cohnii* that showed higher affinity for a TTTT binding motif (34). No such promoter region has yet been identified; however, most sequenced intergenic regions from dinoflagellates are between gene array members. The promoter may be upstream of the entire array. Because dinoflagellate arrays are composed of the same gene, rather than different genes as in trypanosomes, individual gene families could still be regulated at the transcriptional level from this upstream promoter. The implications of these differences in the arrangement of gene arrays in dinoflagellates and trypanosomes have not yet been investigated.

Microarrays and next-generation profiling techniques have also been used to investigate the regulation of gene expression in dinoflagellates. Similarities between trypanosome and dinoflagellate transcript processing and gene structure suggest that dinoflagellates may rely heavily on posttranscriptional mechanisms of gene regulation. Comparative or quantitative gene expression methods measure the transcript abundance within the cells at the time of RNA extraction. Changes in transcript abundance are often attributed to up- or downregulation of gene transcription. However, transcript abundance is dependent on the rate of transcription and on the rate of transcript degradation. When interpreting the results of gene expression analyses in dinoflagellates, it is important to be aware of this dynamic, especially in light of the similarities between dinoflagellate and trypanosome mRNA processing.

Several gene expression profiling studies with dinoflagellates have shown that transcript abundance for a small but significant portion of the transcriptome does respond to various conditions. The first microarray from a dinoflagellate contained 3,500 spotted cDNAs from *Pyrocystis lunula* and was used to identify genes that respond to the circadian clock and oxidative stress (80, 81). About 3% of the genes showed significant changes in transcript abundance in response to a circadian clock and 4% responded to oxidative stress. A microarray containing 4,629 unique genes from *Karenia brevis* showed about 10% of transcripts varying over the diel (24-h) cycle and, similar to *P. lunula*, about 3% appeared to be under circadian control (114). cDNA arrays have been used to try to identify genes involved in toxicity by comparing gene expression levels between toxin-producing and nontoxic strains of species of *Alexandrium minutum* and *K. brevis* (68, 122). Both studies showed small proportions of genes exhibiting significantly different transcript abundance (4% and 7%, respectively), although these arrays probed only a fraction of the transcriptome of these organisms. These results are important for understanding the physiological effects of producing secondary metabolites in these organisms. For example, the nontoxic *K. brevis* showed significantly higher transcript abundance (two- to

Gene amplification: the production of multiple copies of a gene resulting in increased gene expression

threefold increase) for a number of photosystem genes, whereas photosynthesis genes were absent from the set of transcripts with decreased abundance. The relationship between toxin production and photosynthesis is not clear, but this study demonstrated that there are significant physiological differences between toxic and nontoxic strains.

The next few years will see a dramatic expansion in transcriptome data due to the proliferation of RNA-seq transcriptome-profiling methods. The first attempts at global gene expression profiling in dinoflagellates used massively parallel signature sequencing. Two studies profiled gene expression in saxitoxin-producing species of *Alexandrium*, comparing global gene expression profiles under various nutrient conditions such as nutrient-replete conditions, nitrate- and phosphate-limiting conditions, and bacterized xenic cultures (28, 73). These experiments showed that a large proportion of the transcriptome maintained constant mRNA abundance across different conditions, but that a larger proportion of genes (~27%) exhibited variable transcript levels compared to previous microarray experiments. Moustafa et al. (73) identified 40,000 unique signatures, providing an estimate for the number of transcribed genes in *Alexandrium*, and showed that only a handful of transcripts were unique to any particular condition. They identified 56 gene families with more than 100 members, supporting the conclusions of other studies, suggesting that many dinoflagellate genes are encoded in large tandem arrays. The presence of large gene families also suggests that the large number of transcripts in *Alexandrium* might represent a significantly smaller number of unique proteins. The xenic condition had the highest number of unique transcripts (487) and differentially expressed genes, indicating that associated bacteria have a significant impact on the trophic state of *Alexandrium*.

ORGANELLAR GENOMES

Whereas dinoflagellates have some of the largest nuclear genomes of any eukaryote, their organellar genomes are among the smallest in terms of gene number. Only 3 protein-coding genes have been found in the mitochondria, and only 16 genes in the most common plastid, containing the photopigment peridinin. Both organelles must rely on the import of proteins (e.g., polymerases, ribosomal proteins) and tRNAs into the organelle for genome replication and expression of the few genes that remain. Each organelle has evolved novel genome organization and mechanisms of mRNA processing and is a model for understanding genome reduction in mitochondria and plastids.

The Mitochondrial Genome

The biology of dinoflagellate mitochondrial genomes was recently reviewed by Waller & Jackson (115), so only the key features are summarized here. The mitochondrial genome surveys of dinoflagellates have uncovered just three protein-coding genes (*cob*, *cox1*, and *cox3*), two highly fragmented rRNAs, and no tRNAs (44, 48, 74) (**Figure 2*c***). In the basal dinoflagellate *Oxyrrhis*, the *cob* and *cox3* genes are fused, reducing the protein-coding gene count to two, possibly the fewest mitochondria-encoded genes of any organism (106). Apicomplexans also code for the same three protein-coding genes as well as fragmented large and small subunit rRNAs (30). This is in contrast to the more typical arrangement of ciliate mitochondrial genomes, which contain genes for two rRNAs, seven tRNAs, and 43 proteins (17, 88). Ciliate genomes presumably represent the ancestral condition for the more extreme genome configurations within the other alveolates.

Whereas dinoflagellates and apicomplexans share similar mitochondrial gene content, they have different gene arrangements. The apicomplexan mitochondrial genome is linear and compactly arranged on a contiguous 6-kb stretch of DNA. In contrast, the dinoflagellate mitochondrial genes are multicopy and the genome itself is highly fragmented (44, 74, 76). No full-length DNA molecule has been described because large amounts of inverted

repeats and other noncoding DNA hamper sequencing and assembly efforts.

Several unique processes characterize mitochondrial transcription in dinoflagellates. At least one of the mitochondrial genes, *cox3*, is encoded in two genomic elements, which are separately polyadenylated and *trans*-spliced to create one complete mature transcript (44). Transcripts also lack canonical start and stop codons (44, 74, 106). Mitochondrial mRNA is extensively edited, and all but 3 of the 12 possible nucleotide changes have been identified in dinoflagellates (44, 48, 56, 74, 126). No mRNA editing was found in the basal dinoflagellate *Oxyrrhis*, suggesting the species diverged prior to the evolution of editing in the group (106). The function of mRNA editing in dinoflagellates is unclear. However, the majority of edits are to either a G or a C, thereby reducing the overall AT content of mitochondrial transcripts and potentially making them better suited for imported cytosolic tRNAs.

Dinoflagellate Plastids and Their Genomes

About half of all dinoflagellates are photosynthetic (96). The vast majority of these species have a plastid surrounded by three membranes containing the photopigment peridinin, in addition to chlorophylls *a* and *c* (10). Phylogenetic analyses show that the plastids of all chlorophyll *c*–containing algae (dinoflagellates, stramenopiles, haptophytes, and cryptophytes) are monophyletic and derived via endosymbiosis from red algae (8, 39, 84, 86). Plastids in these lineages were thought to be derived from a single secondary endosymbiosis in a common ancestor (i.e., the chromalveolate hypothesis; reviewed in Reference 4). However, nuclear gene phylogenies are not congruent with those from plastid and plastid-targeted genes, suggesting a secondary plastid was acquired in one or more lineages and subsequently transferred to other lineages through tertiary endosymbioses. The common ancestor of dinoflagellates and apicomplexans was likely photosynthetic because the apicomplexans also contain a reduced plastid. The recent discovery of a free-living marine alga, *Chromera velia*, that is sister to the apicomplexan parasites lends strong support to this hypothesis (45, 70).

Most plastid genomes in photosynthetic eukaryotes are single circular molecules derived from the genome of a cyanobacterial endosymbiont and are about 150 kb in length, encoding approximately 100 genes. The plastid genome of peridinin-containing dinoflagellates is remarkable because it has been broken into 2- to 3-kb minicircles, encoding one to several genes and a noncoding core sequence thought to contain the origin of replication (50, 128) (**Figure 2*d***). Recent studies suggest that both replication and transcription of the plastid minicircles take place in a rolling circle fashion (24, 53). Dang & Green (24) proposed that the entire minicircle is transcribed continuously many times around the molecule. The resulting RNAs are cleaved by an endonuclease into long precursor mRNA molecules before finally being processed into a mature mRNA, including polyuridylylation of the 3′ end (117). Thus far, only 16 minicircle-encoded genes have been described (reviewed in Reference 42), making these genomes the smallest plastid genomes known. This small plastid-encoded gene set was confirmed by sequencing a cDNA library constructed by isolating poly(U) mRNAs (117). The genes that remain encoded on the minicircles are 12 core subunits of the four major components of the photosystem (photosystem I and II, cytochrome b_6-f complex, and ATP synthase), two ribosomal proteins, and two conserved hypothetical proteins. The remaining genes required for photosynthesis have been relocated to the nuclear genome (5, 38). It is unclear why such a dramatic reduction of the plastid genome occurred in the dinoflagellates, but it appears to have occurred early in their evolutionary history (96).

The extreme reduction of the plastid genome makes peridinin dinoflagellates a model for understanding the retention of genes in organelles. Organelle-encoded genes are subject to high mutation rates because of the effects of oxygen free radicals, and they

Plastid minicircle: small circular DNA molecules encoding several genes that collectively comprise the plastid genome of peridinin-containing dinoflagellates

Kleptoplast: a temporary plastid retained from prey

LGT: lateral gene transfer

accumulate mutations more quickly because they are encoded on nonrecombining genomes (62). The genes transferred to the nucleus include 15 genes that are encoded on the plastid genome of every other photosynthetic eukaryote (38). The genes that remain can be explained by the colocalization for redox regulation hypothesis, which states that genes for the core components of the photosystem are under selection to remain in the organelle so their products can respond to changes in redox conditions (1). If the genes were encoded in the nucleus, organisms with multiple plastids would not be able to target gene products to a particular organelle. Therefore, the plastid genomes of peridinin-containing dinoflagellates may represent the minimal set of genes that must remain in the organelle of a photosynthetic organism to maintain proper redox regulation.

Nuclear-encoded plastid genes in dinoflagellates have a tripartite N-terminal extension containing a signal peptide, a plastid transit peptide, and a hydrophobic region for targeting and importing into the three-membrane plastid (75). The hydrophobic region is believed to facilitate the insertion of plastid proteins into transport vesicles, which then fuse with the outer membrane of the plastid. These targeting peptides are a striking example of convergent evolution. Targeting peptides with a similar structure have evolved independently in the euglenozoan *Euglena*, which also has a three-membrane plastid acquired independently from green algae (75). Another distinctive aspect of peridinin plastids is the replacement of the typical form I ribulose-1,5-bisphosphate carboxylase-oxygenase (RuBisCO) with a nuclear-encoded form II RuBisCO, which has a significantly lower affinity for CO_2 than the form I enzyme (46, 71, 93). Form II RuBisCo is typically found in proteobacteria growing in high-CO_2, low-O_2 environments. Dinoflagellates may possess novel carbon-concentrating mechanisms to compensate for the enzyme's lower affinity for CO_2 (9). This protein was recently identified in two other alveolates, *Chromera velia* and a related unnamed species, CCMP3115, indicating that form II RuBisCo was acquired in the common ancestor of dinoflagellates and apicomplexans (45) (**Figure 1**).

Dinoflagellates are distinctive among eukaryotes in their propensity for plastid replacement. In addition to the ancestral peridinin plastid, permanent plastids from multiple algal sources have been acquired by some taxa (102). The plastids of *Lepidodinium* spp. are derived from green algae, whereas *Karenia* and *Karlodinium* species have plastids from haptophytes (111, 118). The plastids of several related genera, including *Kryptoperidinium* and *Durinskia* (the dinotoms), are derived from diatoms and are unusual in that the diatom nucleus is retained (23, 43). The plastid genomes of these dinoflagellates with more recently acquired plastids likely have a typical structure (i.e., single circular molecule). This has been confirmed in *Kryptoperidinium* and *Durinskia*, which have circular plastid genomes of 116 kb and 140 kb, respectively (43). Some dinoflagellates, such as *Dinophysis* spp. and *Amphidinium poecilchroum*, are kleptoplastidic, retaining temporary plastids from photosynthetic prey (100). All the currently known plastid replacements have resulted in the loss of characteristics unique to the peridinin plastid (i.e., peridinin, minicircle chromosomes, form II RuBisCo). However, many nuclear-encoded genes from the peridinin-containing ancestor remain in addition to genes derived from the nucleus of the plastid donor (see Endosymbiotic Gene Transfer in the Evolution of Dinoflagellate Genomes, below).

EVOLUTION OF DINOFLAGELLATE GENE CONTENT

Gene duplication and lateral gene transfer (LGT) are two drivers of genome evolution and the evolution of novel traits. LGT has long been recognized to play a major role in the evolution of prokaryotic organisms (e.g., 52, 79, 83). In multicellular eukaryotes, gene duplication plays a major role in the evolution

of new genes because the isolation of germ cells from somatic tissue limits the potential for LGT, although some examples are known (e.g., 27). The relative contribution of these two processes is less understood in microbial eukaryotes, but phylogenomic analyses have revealed an important role for gene transfer in these organisms (3, 16, 49, 72, 77). Dinoflagellates are no exception; LGT from bacteria and other eukaryotes has had a major impact on gene content in this lineage. There are several potential sources for genes acquired through LGT in dinoflagellates. Plastid endosymbiosis has clearly played a major role in the evolution of dinoflagellates and provides the potential for large-scale gene transfer. Dinoflagellates are also prolific grazers of other organisms and many photosynthetic species are mixotrophic. They are also host to many intra- and extracellular bacterial symbionts that could provide genes through LGT. Determining the relative contribution of these potential sources to the evolution of novel traits in dinoflagellates is a major goal of future research. Progress will be facilitated by improvements in the depth and diversity of genome data from dinoflagellates and other microbial eukaryotes.

Endosymbiotic Gene Transfer in the Evolution of Dinoflagellate Genomes

Gene transfer from the endosymbiont to the host nucleus is a critical process in the reduction of an endosymbiont to a plastid. Some genes are retained in the organellar genome, but the majority of genes involved in plastid function have been transferred to the nucleus, a process termed endosymbiotic gene transfer, or EGT (112). The majority of transferred genes are targeted back to the organelle; however, some genes have been co-opted by the host cell to function in novel processes. The primary endosymbiosis, which gave rise to the plastids of the Archaeplastidia (e.g., land plants), resulted in the transfer of a large number of genes from cyanobacteria to eukaryotes. Many of these genes are targeted back to the chloroplast, but some genes function in other cellular processes including disease resistance and protein trafficking (63). During secondary and tertiary endosymbiosis, in which a photosynthetic eukaryote is engulfed by another eukaryote, the nucleus of the endosymbiont is reduced through gene loss and gene transfer to the host. These secondary and tertiary endosymbioses present the opportunity for large-scale gene transfer from one eukaryote to another. The dinoflagellates are remarkable among eukaryotes for their diverse array of plastids acquired through multiple endosymbioses. Phylogenomic analyses have begun to elucidate the role of EGT in dinoflagellate evolution, but the lack of genome-wide data for many dinoflagellates and plastid donors complicates the interpretation of these data.

Endosymbiotic gene transfer (EGT): the transfer of genes from an endosymbiont to the host genome

Analyses of transcriptome data from dinoflagellates have revealed that their genes are derived from many sources, but EGT has likely played a large role in the contribution of genes to the nuclear genome. So far, EGT has contributed genes primarily of plastid function. The evolutionary history of the peridinin-containing plastid is unclear, but it is ultimately derived from a red alga, probably through tertiary endosymbiosis (4). This uncertainty makes determining the evolutionary history of these genes difficult; however, it is clear that plastid genes in peridinin-containing dinoflagellates have multiple origins. Because the other members of the alveolates (ciliates and apicomplexans) are not photosynthetic, the expected sister group for vertically inherited photosynthesis-related genes is the stramenopiles (**Figure 1**). Many nuclear-encoded plastid genes from peridinin dinoflagellates show a sister relationship to stramenopiles, but some cluster with green algae and haptophytes as well (38, 116). Dinoflagellates with more recently acquired plastids have a large number of genes from the donor of the most recent organelle; however, they have genes from other algal sources as well. *Karenia* and *Karlodinium*, genera that contain a haptophyte plastid, have plastid-related genes retained from a peridinin-containing ancestor and those acquired from other algae, in addition to genes contributed by the haptophyte

symbiont (78, 85, 123). Likewise, *Lepidodinium viride*, which has a plastid from prasinophyte green algae, contains nuclear-encoded plastid genes transferred from the symbiont, but it also has genes from other sources (66, 109).

Following endosymbiosis, the transfer of large numbers of plastid-targeted genes from the endosymbiont is expected, since these genes are already adapted to function in the newly acquired organelle. The retention of genes from a peridinin-containing ancestor in species with more recently acquired plastids has been interpreted as evidence for plastid replacement, rather than complete plastid loss and regain (66, 78, 85). However, the discovery of plastid-related genes in a heterotrophic dinoflagellate indicates that some of these genes could have been retained through a heterotrophic period between plastid loss and acquisition of a new organelle (98). The kleptoplastidic dinoflagellate *Dinophysis acuminata*, which has temporary plastids retained from prey, has a small number of nuclear-encoded plastid genes derived from multiple algal lineages (120). This organism is notable because it does not contain a permanent plastid, yet it encodes genes for plastid function in its nucleus. Organisms such as *Dinophysis* may provide insights into the process of endosymbiosis through prey organelle retention.

Genes contributed by other algal lineages could be the result of sporadic transfers of individual genes, perhaps from prey, or they may have been acquired through cryptic endosymbioses. Distinguishing between these two possibilities has, thus far, been difficult in the dinoflagellates because of incomplete genome datasets and the limited diversity of data from potential plastid donors. Current phylogenetic analyses lack the resolution to determine whether significant numbers of genes originate from the same donor, as would be expected from EGT. Recently, a cryptic endosymbiosis was suggested to explain a large number of genes in diatoms (17% of the proteome) that appear to be derived from green algae (72). This conclusion was possible because of the availability of complete genome sequences for the diatoms (*Thalassiosira* and *Phaeodactylum*) as well as genome data from a diverse assemblage of green algae, which allowed the identification of a large number of diatom genes that branched at the same position within the green algal phylogeny. In the near future, the availability of genome-wide data for dinoflagellates and diverse genome data from potential donor groups such as haptophytes, cryptophytes, and stramenopiles will help to provide a clearer picture of dinoflagellate gene origin.

Lateral Gene Transfer from Bacteria

Phylogenomic analyses are revealing an increasing number of genes transferred from bacteria into the genomes of microbial eukaryotes (3, 49, 77). This process is responsible for some of the most distinctive features of dinoflagellate biology. The histone-like proteins, which play a role in the unique chromosome structure of dinoflagellates, and form II RuBisCO were both acquired from proteobacteria (36, 45, 71, 121). In another case, *aroB* and an *O*-methyl transferase gene from cyanobacteria were transferred into dinoflagellates between the divergence of *Perkinsus* and *Oxyrrhis* (116). These two genes fused and formed a novel plastid-targeted gene not found in any other eukaryotic lineage. Bacteria have likely contributed many more genes to the dinoflagellates, and future whole-genome phylogenomic analyses will further clarify the prokaryotic contribution to the evolution of novel traits in these enigmatic eukaryotes.

CONCLUDING REMARKS

Next-generation sequencing technologies and phylogenomic tools have set the stage for genomics to revolutionize our understanding of dinoflagellate genomes, ecology, and evolution. The dramatically decreasing cost of producing genome data has put tools in the hands of individual researchers that were only available to large consortiums of scientists a few years ago. A dinoflagellate genome sequence is no longer out of reach and will likely be completed within the next few years, if not months, and

followed by many others. As an alternative to genome sequencing, transcriptome sequencing provides a cost-effective way to determine the gene complement of these organisms without the effort involved in assembling and annotating a genome. These methods will be critical to expanding the diversity of genomic data for dinoflagellates and other microbial eukaryotes that will be critical to phylogenomic studies and resolving the dinoflagellate phylogenetic tree. Transcriptomics also has been, and will continue to be, a critical tool for determining gene function in the dinoflagellates, which lack a genetic system for experimentation. Future challenges for scientists will be to develop methods to analyze upcoming data to clarify the important roles of dinoflagellates in ocean ecology as well as their interactions with other organisms, and to understand the processes that have resulted in the evolution of so many amazingly different solutions to genome structure, organization, and regulation in the dinoflagellates.

SUMMARY POINTS

1. The dinoflagellate nuclear genome has several distinctive characteristics including permanently condensed liquid-crystalline chromosomes that lack nucleosomes comprising their large genomes.
2. Dinoflagellates have large numbers of genes encoded in tandem gene arrays, suggesting that gene amplification has been an important process in these organisms.
3. 5′ *trans*-spliced leader addition during mRNA processing may indicate that dinoflagellates rely heavily on posttranscriptional regulation of gene expression.
4. Quantitative gene expression analyses show that a portion of the transcriptome does respond to changing conditions.
5. LGT has played a major role in the evolution of dinoflagellate gene content.
6. EGT has been a major mechanism of genome evolution.
7. Genes transferred from bacteria are involved in some of the most distinctive traits in dinoflagellates.

FUTURE ISSUES

1. When did novel dinoflagellate traits (e.g., dinokaryon, *trans*-splicing, and plastid minicircles) evolve?
2. Do the extrachromosomal loops encode gene under active transcription?
3. What are the roles of histones and HLPs in chromosome structure?
4. What are the mechanisms and relative importance of transcriptional and posttranscriptional gene regulation?
5. How have LGT and EGT contributed to the evolution of novel dinoflagellate traits?

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

1. Allen JF. 2003. The function of genomes in bioenergetic organelles. *Phil. Trans. R. Soc. B* 358:19–37
2. Allen JR, Roberts TM, Loeblich AR, Klotz LC. 1975. Characterization of DNA from dinoflagellate *Crypthecodinium cohnii* and implications for nuclear organization. *Cell* 6:161–69
3. Andersson JO, Sarchfield SW, Roger AJ. 2005. Gene transfers from nanoarchaeota to an ancestor of diplomonads and parabasalids. *Mol. Biol. Evol.* 22:85–90
4. Archibald JM. 2009. The puzzle of plastid evolution. *Curr. Biol.* 19:81–88
5. Bachvaroff TR, Concepcion GT, Rogers CR, Herman EM, Delwiche CF. 2004. Dinoflagellate expressed sequence tag data indicate massive transfer of chloroplast genes to the nuclear genome. *Protist* 155:65–78

6. Shows tandem gene array arrangement of highly expressed genes, intron distribution, and a putative polyA signal.

6. **Bachvaroff TR, Place AR. 2008. From stop to start: tandem gene arrangement, copy number and *trans*-splicing sites in the dinoflagellate *Amphidinium carterae*. *PLoS One* 3:e2929**
7. Bachvaroff TR, Place AR, Coats DW. 2009. Expressed sequence tags from *Amoebophrya* sp. infecting *Karlodinium veneficum*: comparing host and parasite sequences. *J. Eukaryot. Microbiol.* 56:531–41
8. Bachvaroff TR, Puerta MVS, Delwiche CF. 2005. Chlorophyll *c*–containing plastid relationships based on analyses of a multigene data set with all four chromalveolate lineages. *Mol. Biol. Evol.* 22:1772–82
9. Badger MR, Bek EJ. 2008. Multiple Rubisco forms in proteobacteria: their functional significance in relation to CO_2 acquisition by the CBB cycle. *J. Exp. Bot.* 59:1525–41
10. Barsanti L, Gualtieri P. 2006. *Algae Anatomy, Biochemistry, and Biotechnology*. New York: CRC Press
11. Bhattacharya D, Yoon HS, Hedges SB, Hackett JD. 2009. Eukaryotes (Eukaryota). In *The Timetree of Life*, ed. SB Hedges, S Kumar, pp. 116–20. Oxford, UK: Oxford Univ. Press
12. Bodansky S, Mintz LB, Holmes DS. 1979. Mesokaryote *Gyrodinium cohnii* lacks nucleosomes. *Biochem. Biophys. Res. Commun.* 88:1329–36
13. Bonetta L. 2010. Whole-genome sequencing breaks the cost barrier. *Cell* 141:917–19
14. Boorstein RJ, Chiu LN, Teebor GW. 1989. Phylogenetic evidence of a role for 5-hydroxymethyluracil-DNA glycosylase in the maintenance of 5-methylcytosine in DNA. *Nucleic Acids Res.* 17:7653–61
15. Bouligand Y, Norris V. 2001. Chromosome separation and segregation in dinoflagellates and bacteria may depend on liquid crystalline states. *Biochimie* 83:187–92
16. Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, et al. 2008. The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 456:239–44
17. Burger G, Zhu Y, Littlejohn TG, Greenwood SJ, Schnare MN, et al. 2000. Complete sequence of the mitochondrial genome of *Tetrahymena pyriformis* and comparison with *Paramecium aurelia* mitochondrial DNA. *J. Mol. Biol.* 297:365–80
18. Burki F, Shalchian-Tabrizi K, Minge M, Skjæveland A, Nikolaev SI, et al. 2007. Phylogenomics reshuffles the eukaryotic supergroups. *PLoS One* 8:e790
19. Burki F, Shalchian-Tabrizi K, Pawlowski J. 2008. Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes. *Biol. Lett.* 4:366–69
20. Cachon J, Cachon M. 1987. Parasitic dinoflagellates. In *The Biology of the Dinoflagellates*, ed. FJR Taylor, pp. 571–610. Oxford, UK: Blackwell Sci.
21. Cavalier-Smith T. 1991. Cell diversification in heterotrophic flagellates. In *The Biology of Free-living Heterotrophic Flagellates*, ed. DJ Patterson, J Larson, pp. 113–31. Oxford: Clarendon Press
22. Chan YH, Wong JT. 2007. Concentration-dependent organization of DNA by the dinoflagellate histone-like protein HCc3. *Nucleic Acids Res.* 35:2573–83

23. Reviews concentration-dependent compaction of DNA by histone-like proteins.

23. **Chesnick J, Kooistra W, Wellbrock U, Medlin L. 1997. Ribosomal RNA analysis indicates a benthic pennate diatom ancestry for the endosymbionts of the dinoflagellates *Peridinium foliaceum* and *Peridinium balticum* (Pyrrhophyta). *J. Eukaryot. Microbiol.* 44:314–20**

24. Dang YK, Green BR. 2010. Long transcripts from dinoflagellate chloroplast minicircles suggest "rolling circle" transcription. *J. Biol. Chem.* 285:5196–203
25. Davies W, Jakobsen KS, Nordby O. 1988. Characterization of DNA from the dinoflagellate *Woloszynskia bostoniensis*. *J. Protozool.* 35:418–22
26. Dodge JD. 1965. Chromosome structure in the dinoflagellates and the problem of the mesocaryotic cell. *Excerpta Med. Int. Congr. Ser.* 91:339–45
27. Dunning Hotopp JC, Clark ME, Oliveira DC, Foster JM, Fischer P, et al. 2007. Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science* 317:1753–56
28. Erdner DL, Anderson DM. 2006. Global transcriptional profiling of the toxic dinoflagellate *Alexandrium fundyense* using massively parallel signature sequencing. *BMC Genomics* 7:88
29. Fast NM, Xue LR, Bingham S, Keeling PJ. 2002. Re-examining alveolate evolution using multiple protein molecular phylogenies. *J. Eukaryot. Microbiol.* 49:30–37
30. Feagin JE. 1992. The 6-Kb element of *Plasmodium falciparum* encodes mitochondrial cytochrome genes. *Mol. Biochem. Parasitol.* 52:145–48
31. Fensome RA, Taylor FJ, Norris G, Sarjeant WAS, Warton DI, Williams GL. 1993. A classification of living and fossil dinoflagellates. *Am. Mus. Nat. Hist., Micropaleontol. Spec. Publ.* 7, 351 pp.
32. Gautier A, Michel-Salamin L, Tosi-Couture E, McDowall AW, Dubochet J. 1986. Electron microscopy of the chromosomes of dinoflagellates in situ: confirmation of Bouligand's liquid crystal hypothesis. *J. Ultrastruct. Mol. Struct. Res.* 97:10–30
33. Graham L, Wilcox L. 2000. *Algae*. Upper Saddle River, NJ: Prentice-Hall
34. Guillebault D, Sasorith S, Derelle E, Wurtz JM, Lozano JC, et al. 2002. A new class of transcription initiation factors, intermediate between TATA box-binding proteins (TBPs) and TBP-like factors (TLFs), is present in the marine unicellular organism, the dinoflagellate *Crypthecodinium cohnii*. *J. Biol. Chem.* 277:40881–86
35. Gunderson JH, Goss SH, Coats DW. 1999. The phylogenetic position of *Amoebophrya* sp. infecting *Gymnodinium sanguineum*. *J. Eukaryot. Microbiol.* 46:194–97
36. Hackett JD, Scheetz TE, Yoon HS, Soares MB, Bonaldo MF, et al. 2005. Insights into a dinoflagellate genome through expressed sequence tag analysis. *BMC Genomics* 6:80
37. Hackett JD, Yoon HS, Butterfield NJ, Sanderson MJ, Bhattacharya D. 2007. Plastid endosymbiosis: sources and timing of the major events. In *Evolution of Primary Producers in the Sea*, ed. PG Falkowski, AH Knoll, pp. 109–32. San Diego, CA: Academic
38. Hackett JD, Yoon HS, Soares MB, Bonaldo MF, Casavant TL, et al. 2004. Migration of the plastid genome to the nucleus in a peridinin dinoflagellate. *Curr. Biol.* 14:213–18
39. Harper JT, Keeling PJ. 2003. Nucleus-encoded, plastid-targeted glyceraldehyde-3-phosphate dehydrogenase (GAPDH) indicates a single origin for chromalveolate plastids. *Mol. Biol. Evol.* 20:1730–35
40. Hoet PP, Coene MM, Cocito CG. 1992. Replication cycle of *Bacillus subtilis* hydroxymethyluracil-containing phages. *Annu. Rev. Microbiol.* 46:95–116
41. Hoppenrath M, Leander BS. 2010. Dinoflagellate phylogeny as inferred from heat shock protein 90 and ribosomal gene sequences. *PLoS One* 5:e13220
42. Howe CJ, Nisbet RER, Barbrook AC. 2008. The remarkable chloroplast genome of dinoflagellates. *J. Exp. Bot.* 59:1035–45
43. Imanian B, Pombert JF, Keeling PJ. 2010. The complete plastid genomes of the two 'dinotoms' *Durinskia baltica* and *Kryptoperidinium foliaceum*. *PLoS One* 5:e10711
44. Jackson CJ, Norman JE, Schnare MN, Gray MW, Keeling PJ, Waller RF. 2007. Broad genomic and transcriptional analysis reveals a highly derived genome in dinoflagellate mitochondria. *BMC Biol.* 5:41
45. Janouskovec J, Horak A, Obornik M, Lukes J, Keeling PJ. 2010. A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proc. Natl. Acad. Sci. USA* 107:10949–54
46. Jordan DB, Ogren WL. 1981. A sensitive assay procedure for simultaneous determination of ribulose-1,5-bisphosphate carboxylase and oxygenase activities. *Plant Physiol.* 67:237–45
47. Joseph SJ, Fernandez-Robledo JA, Gardner MJ, El-Sayed NM, Kuo CH, et al. 2010. The alveolate *Perkinsus marinus*: biological insights from EST gene discovery. *BMC Genomics* 11:228

48. Kamikawa R, Nishimura H, Sako Y. 2009. Analysis of the mitochondrial genome, transcripts, and electron transport activity in the dinoflagellate *Alexandrium catenella* (Gonyaulacales, Dinophyceae). *Phycol. Res.* 57:1–11
49. Keeling PJ, Palmer JD. 2008. Horizontal gene transfer in eukaryotic evolution. *Nat. Rev. Genet.* 9:605–18
50. Koumandou VL, Nisbet RER, Barbrook AC, Howe CJ. 2004. Dinoflagellate chloroplasts—Where have all the genes gone? *Trends Genet.* 20:261–67
51. LaJeunesse TC, Lambert G, Andersen RA, Coffroth MA, Galbraith DW. 2005. *Symbiodinium* (Pyrrhophyta) genome sizes (DNA content) are smallest among dinoflagellates. *J. Phycol.* 41:880–86
52. Lerat E, Daubin V, Ochman H, Moran NA. 2005. Evolutionary origins of genomic repertoires in bacteria. *PLoS Biol.* 3:e130
53. Leung SK, Wong JTY. 2009. The replication of plastid minicircles involves rolling circle intermediates. *Nucleic Acids Res.* 37:1991–2002
54. Levine N. 1970. Taxonomy of the Sporozoa. *J. Parasitol.* 56:208–9
55. Lidie KB, van Dolah FM. 2007. Spliced leader RNA-mediated *trans*-splicing in a dinoflagellate, *Karenia brevis*. *J. Eukaryot. Microbiol.* 54:427–35
56. Lin S, Zhang H, Spencer DF, Norman JE, Gray MW. 2002. Widespread and extensive editing of mitochondrial mRNAS in dinoflagellates. *J. Mol. Biol.* 320:727–39
57. Lin S, Zhang H, Zhuang Y, Tran B, Gill J. 2010. Spliced leader-based metatranscriptomic analyses lead to recognition of hidden genomic features in dinoflagellates. *Proc. Natl. Acad. Sci. USA* 107:20033–38
58. Liu LY, Hastings JW. 2006. Novel and rapidly diverging intergenic sequences between tandem repeats of the luciferase genes in seven dinoflagellate species. *J. Phycol.* 42:96–103
59. Livolant F, Bouligand Y. 1978. New observations on twisted arrangement of dinoflagellate chromosomes. *Chromosoma* 68:21–44
60. Lopez-Garcia P, Rodriguez-Valera F, Pedros-Alio C, Moreira D. 2001. Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* 409:603–7
61. Lowe CD, Keeling PJ, Martin LE, Slamovits CH, Watts PC, Montagnes DJS. 2011. Who is *Oxyrrhis marina*? Morphological and phylogenetic studies on an unusual dinoflagellate. *J. Plankton Res.* 33:591–602
62. Martin W, Herrmann RG. 1998. Gene transfer from organelles to the nucleus: how much, what happens, and why? *Plant Physiol.* 118:9–17
63. Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, et al. 2002. Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci. USA* 99:12246–51
64. McEwan M, Humayun R, Slamovits CH, Keeling PJ. 2008. Nuclear genome sequence survey of the dinoflagellate *Heterocapsa triquetra*. *J. Eukaryot. Microbiol.* 55:530–35
65. Meng FW, Zhou CM, Yin LM, Chen ZL, Yuan XL. 2005. The oldest known dinoflagellates: morphological and molecular evidence from Mesoproterozoic rocks at Yongji, Shanxi Province. *Chinese Sci. Bull.* 50:1230–34
66. Minge MA, Shalchian-Tabrizi K, Torresen OK, Takishita K, Probert I, et al. 2010. A phylogenetic mosaic plastid proteome and unusual plastid-targeting signals in the green-colored dinoflagellate *Lepidodinium chlorophorum*. *BMC Evol. Biol.* 10:191
67. Moestrup O, Daugbjerg N. 2007. On dinoflagellate phylogeny and classification. In *Unraveling the Algae: The Past, Present, and Future of Algal Systematics*, ed. J Brodie, J Lewis, pp. 215–30. New York: CRC Press
68. Monroe EA, Johnson JG, Wang ZH, Pierce RK, Van Dolah FM. 2010. Characterization and expression of nuclear-encoded polyketide synthases in the brevetoxin-producing dinoflagellate *Karenia brevis*. *J. Phycol.* 46:541–52
69. Moon-van der Staay SY, De Wachter R, Vaulot D. 2001. Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* 409:607–10

51. The most recent and comprehensive survey of dinoflagellate genome size estimates.

70. Moore RB, Obornik M, Janouskovec J, Chrudimsky T, Vancova M, et al. 2008. A photosynthetic alveolate closely related to apicomplexan parasites. *Nature* 451:959–63
71. Morse D, Salois P, Markovic P, Hastings J. 1995. A nuclear-encoded form 1RuBisCO in dinoflagellates. *Science* 268:1622–24
72. Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D. 2009. Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science* 324:1724–26
73. Moustafa A, Evans AN, Kulis DM, Hackett JD, Erdner DL, et al. 2010. Transcriptome profiling of a toxic dinoflagellate reveals a gene-rich protist and a potential impact on gene expression due to bacterial presence. *PLoS One* 5:e9688
74. Nash EA, Barbrook AC, Edwards-Stuart RK, Bernhardt K, Howe CJ, Nisbet RER. 2007. Organization of the mitochondrial genome in the dinoflagellate *Amphidinium carterae*. *Mol. Biol. Evol.* 24:1528–36
75. Nassoury N, Cappadocia M, Morse D. 2003. Plastid ultrastructure defines the protein import pathway in dinoflagellates. *J. Cell Sci.* 116:2867–74
76. Norman JE, Gray MW. 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.* 53:351–63
77. Nosenko T, Bhattacharya D. 2007. Horizontal gene transfer in chromalveolates. *BMC Evol. Biol.* 7:173
78. Nosenko T, Lidie KL, Van Dolah FM, Lindquist E, Cheng JF, Bhattacharya D. 2006. Chimeric plastid proteome in the Florida "red tide" dinoflagellate *Karenia brevis*. *Mol. Biol. Evol.* 23:2026–38
79. Ochman H, Lawrence JG, Groisman EA. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* 405:299–304
80. Okamoto OK, Hastings JW. 2003. Genome-wide analysis of redox-regulated genes in a dinoflagellate. *Gene* 321:73–81
81. Okamoto OK, Hastings JW. 2003. Novel dinoflagellate clock-related genes identified through microarray analysis. *J. Phycol.* 39:519–26
82. Ouellette M, Papadopoulou B. 2009. Coordinated gene expression by post-transcriptional regulons in African trypanosomes. *J. Biol.* 8:100
83. Pal C, Papp B, Lercher MJ. 2005. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nat. Genet.* 37:1372–75
84. Patron NJ, Rogers MB, Keeling PJ. 2004. Gene replacement of fructose-1,6-bisphosphate aldolase supports the hypothesis of a single photosynthetic ancestor of chromalveolates. *Eukaryot. Cell* 3:1169–75
85. Patron NJ, Waller RF, Keeling PJ. 2006. A tertiary plastid uses genes from two endosymbionts. *J. Mol. Biol.* 357:1373–82
86. Petersen J, Teich R, Brinkmann H, Cerff R. 2006. A "green" phosphoribulokinase in complex algae with red plastids: evidence for a single secondary endosymbiosis leading to haptophytes, cryptophytes, heterokonts, and dinoflagellates. *J. Mol. Evol.* 62:143–57
87. Prescott DM. 1994. The DNA of ciliated protozoa. *Microbiol. Rev.* 58:233–67
88. Pritchard AE, Seilhamer JJ, Mahalingam R, Sable CL, Venuti SE, Cummings DJ. 1990. Nucleotide sequence of the mitochondrial genome of *Paramecium*. *Nucleic Acids Res.* 18:173–80
89. Rae PMM. 1973. 5-Hydroxymethyluracil in DNA of a dinoflagellate. *Proc. Natl. Acad. Sci. USA* 70:1141–45
90. Raikov I. 1996. Nuclei of ciliates. In *Ciliates Cells as Organisms*, ed. K Hausmann, P Bradbury, pp. 221–42. Stuttgart, Ger.: Gustav Fischer
91. Reece KS, Siddall ME, Burreson EM, Graves JE. 1997. Phylogenetic analysis of *Perkinsus* based on actin gene sequences. *J. Parasitol.* 83:417–23
92. Rizzo PJ. 1991. The enigma of the dinoflagellate chromosome. *J. Protozool.* 38:246–52
93. Rowan R, Whitney SM, Fowler A, Yellowlees D. 1996. Rubisco in marine symbiotic dinoflagellates: form II enzymes in eukaryotic oxygenic phototrophs encoded by a nuclear multigene family. *Plant Cell* 8:539–53

73. Estimates gene number and gene family sizes, few condition-specific transcripts, and the influence of bacteria on dinoflagellate gene expression.

74. Comprehensive characterization of the mitochondrial genome from a dinoflagellate (for a review on dinoflagellate mitochondrial genomes see Reference 118).

75. Describes plastid targeting signals in peridinin dinoflagellates and shows their similarity to those signals from *Euglena*.

78. Dinoflagellates with haptophyte plastids retain genes from peridinin-containing ancestors (also shown in Reference 85) and other algae.

94. Sala-Rovira M, Geraud ML, Caput D, Jacques F, Soyer-Gobillard MO, et al. 1991. Molecular cloning and immunolocalization of two variants of the major basic nuclear protein (HCc) from the histone-less eukaryote *Crypthecodinium cohnii* (Pyrrhophyta). *Chromosoma* 100:510–18
95. Saldarriaga JF, McEwan ML, Fast NM, Taylor FJ, Keeling PJ. 2003. Multiple protein phylogenies show that *Oxyrrhis marina* and *Perkinsus marinus* are early branches of the dinoflagellate lineage. *Int. J. Syst. Evol. Microbiol.* 53:355–65
96. Saldarriaga JF, Taylor FJR, Keeling PJ, Cavalier-Smith T. 2001. Dinoflagellate nuclear SSU rRNA phylogeny suggests multiple plastid losses and replacements. *J. Mol. Evol.* 53:204–13
97. Saldarriaga JF, Taylor FJRM, Cavalier-Smith T, Menden-Deuer S, Keeling PJ. 2004. Molecular data and the evolutionary history of dinoflagellates. *Eur. J. Protistol.* 40:85–111
98. Sanchez-Puerta MV, Lippmeier JC, Apt KE, Delwiche CF. 2007. Plastid genes in a non-photosynthetic dinoflagellate. *Protist* 158:105–17
99. Saunders GW, Hill DRA, Sexton JP, Andersen RA. 1997. Small-subunit ribosomal RNA sequences from selected dinoflagellates: testing classical evolutionary hypotheses with molecular systematic methods. *Plant Syst. Evol.*: 237–59
100. Schnepf E. 1993. From prey via endosymbiont to plastid: comparative studies in dinoflagellates. In *Origins of Plastids*, ed. RA Lewin, pp. 53–76. New York: Chapman & Hall
101. Seo KS, Fritz L. 2006. Karyology of a marine non-motile dinoflagellate, *Pyrocystis lunula*. *Hydrobiologia* 563:289–96
102. Shalchian-Tabrizi K, Minge MA, Cavalier-Smith T, Nedreklepp JM, Klaveness D, Jakobsen KS. 2006. Combined heat shock protein 90 and ribosomal RNA sequence phylogeny supports multiple replacements of dinoflagellate plastids. *J. Eukaryot. Microbiol.* 53:217–24
103. Sigee DC. 1983. Structural DNA and genetically active DNA in dinoflagellate chromosomes. *Biosystems* 16:203–10
104. Sigee DC. 1986. The dinoflagellate chromosome. *Adv. Bot. Res.* 12:205–60
105. Skovgaard A, Meneses I, Angelico MM. 2009. Identifying the lethal fish egg parasite *Ichthyodinium chabelardi* as a member of marine alveolate group I. *Environ. Microbiol.* 11:2030–41
106. Slamovits CH, Saidarriaga JF, Larocque A, Keeling PJ. 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.* 372:356–68
107. Spector DL. 1984. Dinoflagellate nuclei. In *Dinoflagellates*, ed. DL Spector, pp. 107–47. Orlando: Academic
108. Sutton RE, Boothroyd JC. 1986. Evidence for *trans* splicing in trypanosomes. *Cell* 47:527–35
109. Takishita K, Kawachi M, Noel M, Matsumoto T, Kakizoe N, et al. 2008. Origins of plastids and glyceraldehyde-3-phosphate dehydrogenase genes in the green-colored dinoflagellate *Lepidodinium chlorophorum*. *Gene* 410:26–36
110. Teles-Grilo ML, Duarte SM, Tato-Costa J, Gaspar-Maia A, Oliveira C, et al. 2007. Molecular karyotype analysis of *Perkinsus atlanticus* (phylum Perkinsozoa) by pulsed field gel electrophoresis. *Eur. J. Protistol.* 43:315–18
111. Tengs T, Dahlberg OJ, Shalchian-Tabrizi K, Klaveness D, Rudi K, et al. 2000. Phylogenetic analyses indicate that the 19′hexanoyloxy-fucoxanthin-containing dinoflagellates have tertiary plastids of haptophyte origin. *Mol. Biol. Evol.* 17:718–29
112. Timmis JN, Ayliffe MA, Huang CY, Martin W. 2004. Endosymbiotic gene transfer: Organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* 5:123–35
113. Tourancheau AB, Tsao N, Klobutcher LA, Pearlman RE, Adoutte A. 1995. Genetic code deviations in the ciliates: evidence for multiple and independent events. *EMBO J.* 14:3262–67
114. Van Dolah FM, Lidie KB, Morey JS, Brunelle SA, Ryan JC, et al. 2007. Microarray analysis of diurnal- and circadian-regulated genes in the florida red-tide dinoflagellate *Karenia brevis* (Dinophyceae). *J. Phycol.* 43:741–52
115. Waller RF, Jackson CJ. 2009. Dinoflagellate mitochondrial genomes: stretching the rules of molecular biology. *Bioessays* 31:237–45

116. Waller RF, Slamovits CH, Keeling PJ. 2006. Lateral gene transfer of a multigene region from cyanobacteria to dinoflagellates resulting in a novel plastid-targeted fusion protein. *Mol. Biol. Evol.* 23:1437–43

117. Wang Y, Morse D. 2006. Rampant polyuridylylation of plastid gene transcripts in the dinoflagellate *Lingulodinium*. *Nucleic Acids Res.* 34:613–19

118. Watanabe M, Takeda Y, Sasa T, Inouye I, Suda S, et al. 1987. A green dinoflagellate with chlorophylls *a* and *b*: morphology, fine structure, of the chloroplast and chlorophyll composition. *J. Phycol.* 23:382–89

119. Wilson RJM, Williamson DH. 1997. Extrachromosomal DNA in the Apicomplexa. *Microbiol. Mol. Biol. Rev.* 61:1–16

120. Wisecaver JH, Hackett JD. 2010. Transcriptome analysis reveals nuclear-encoded proteins for the maintenance of temporary plastids in the dinoflagellate *Dinophysis acuminata*. *BMC Genomics* 11:366

121. Wong JTY, New DC, Wong JCW, Hung VKL. 2003. Histone-like proteins of the dinoflagellate *Crypthecodinium cohnii* have homologies to bacterial DNA-binding proteins. *Eukaryot. Cell* 2:646–50

122. Yang I, John U, Beszteri S, Glockner G, Krock B, et al. 2010. Comparative gene expression in toxic versus non-toxic strains of the marine dinoflagellate *Alexandrium minutum*. *BMC Genomics* 11:248

123. Yoon HS, Hackett JD, Van Dolah FM, Nosenko T, Lidie KL, Bhattacharya D. 2005. Tertiary endosymbiosis-driven genome evolution in dinoflagellate algae. *Mol. Biol. Evol.* 22:1299–308

124. Zhang H, Bhattacharya D, Lin S. 2007. A three-gene dinoflagellate phylogeny suggests monophyly of Prorocentrales and a basal position for *Amphidinium* and *Heterocapsa*. *J. Mol. Evol.* 65:463–74

125. Zhang H, Hou Y, Miranda L, Campbell DA, Sturm NR, et al. 2007. Spliced leader RNA *trans*-splicing in dinoflagellates. *Proc. Natl. Acad. Sci. USA* 104:4618–23

126. Zhang H, Hou YB, Lin SJ. 2006. Isolation and characterization of proliferating cell nuclear antigen from the dinoflagellate *Pfiesteria piscicida*. *J. Eukaryot. Microbiol.* 53:142–50

127. Zhang H, Lin S. 2008. mRNA editing and spliced-leader RNA *trans*-splicing groups *Oxyrrhis*, *Noctiluca*, *Heterocapsa*, and *Amphidinium* as basal lineages of dinoflagellates. *J. Phycol.* 44:703–11

128. Zhang ZD, Green BR, Cavalier-Smith T. 1999. Single gene circles in dinoflagellate chloroplast genomes. *Nature* 400:155–59

117. Shows plastid transcripts are polyuridylylated.

120. Reviews plastid-targeted genes in a dinoflagellate with temporary plastids retained from prey.

125. Discusses the addition of 5′ *trans*-spliced leaders during mRNA processing in dinoflagellates (see also Reference 55).